

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DIAGNÓSTICO DE FALHAS EM PROCESSOS INDUSTRIAIS
USANDO CLASSIFICADORES LOCAIS AVALIADOS COM
DIFERENTES CARACTERÍSTICAS**

DISSERTAÇÃO DE MESTRADO

GUSTAVO BOINA SANTI

VITÓRIA

2018

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DIAGNÓSTICO DE FALHAS EM PROCESSOS INDUSTRIAIS
USANDO CLASSIFICADORES LOCAIS AVALIADOS COM
DIFERENTES CARACTERÍSTICAS**

GUSTAVO BOINA SANTI

Dissertação de Mestrado apresentada como
requisito parcial para obtenção do título de
Mestre em Engenharia Elétrica.

Orientador:

Prof. Dr. Celso José Munaro.

VITÓRIA

2018

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial Tecnológica, Universidade Federal do
Espírito Santo, ES, Brasil)

S235d Santi, Gustavo Boina, 1992-
Diagnóstico de falhas em processos industriais usando
classificadores locais avaliados com diferentes
características / Gustavo Boina Santi. – 2018.
90 f. : il.

Orientador: Celso José Munaro.
Dissertação (Mestrado em Engenharia Elétrica) –
Universidade Federal do Espírito Santo, Centro
Tecnológico.

1. Análise multivariada. 2. Localização de falhas. 3.
Análise por agrupamento. 4. Classificadores (Linguística).
5. Seleção de características (Computação). 6.
Visualização de dados. I. Munaro, Celso José. II.
Universidade Federal do Espírito Santo. Centro
Tecnológico. III. Título.

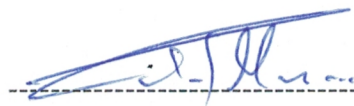
CDU: 621.3

GUSTAVO BOINA SANTI

**DIAGNÓSTICO DE FALHAS EM PROCESSOS INDUSTRIAIS
USANDO CLASSIFICADORES LOCAIS AVALIADOS COM
DIFERENTES CARACTERÍSTICAS**

Dissertação submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Engenharia Elétrica.

COMISSÃO EXAMINADORA



Prof. Dr. Celso José Munaro
Universidade Federal do Espírito Santo
Orientador



Prof. Dr. Reinaldo Martinez Palhares
Universidade Federal de Minas Gerais



Prof. Dr. Francisco de Assis Boldt
Instituto Federal do Espírito Santo

“The desire for knowledge shapes a man”

(Patrick Rothfuss)

Agradecimentos

Aos meus pais por sempre terem me apoiado em minhas escolhas e estarem ao meu lado em todos os momentos da minha vida. Aos amigos de curso pela ajuda nas horas difíceis e aos colegas do Laboratório de Controle e Instrumentação (LCI) por todo apoio, disposição e auxílio que me deram no decorrer desse trabalho, são eles: Arthur Almeida Assis, Diego Calegario, Débora Martins, Felipe Machado Lobo, Heitor Guzzo Faria, Marcos Vinicius Gomes Cypriano, Óscar Becerra, Victor Rebli Nunes e Vinícius Belmuds Vasconcelos. Ao orientador, Dr. Celso José Munaro, pelo conhecimento e direcionamento dados. Ao Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da UFES por todo apoio dado e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.

Resumo

A ocorrência de falhas em processos industriais acarreta a operação indevida de equipamentos, o que é um risco para a segurança dos operadores e reduzem os ganhos. Nesse contexto a realização de um diagnóstico de falhas preciso é de fundamental importância. Nesse trabalho é apresentada uma metodologia para diagnóstico de falhas em processos industriais. São utilizados classificadores juntamente com técnicas de análise multivariada de dados para diagnóstico das falhas. Para melhorar o diagnóstico, é proposto o agrupamento de falhas que possuam influência em variáveis semelhantes. Dessa forma, ao invés de um único classificador, é proposto o uso de múltiplos classificadores locais. Utilizar um classificador para todas as falhas pode tornar a tarefa de classificação mais complexa e reduzir a acurácia do diagnóstico, enquanto classificadores locais podem ser mais simples e terem maior poder de discriminação entre as falhas. Para simplificar os dados a serem analisados pelos classificadores, a abordagem proposta faz o uso de extração de características para que seja analisado o comportamento do processo durante o avanço de uma falha. As atividades são realizadas no ambiente MATLAB e a metodologia é aplicada a três estudos de caso: classificação de séries temporais disponíveis em bases de dados consolidadas na literatura, simulações da planta Tennessee Eastman e simulações de um reator tanque agitado contínuo.

Abstract

Faults in industrial process lead to equipment malfunction, which can cause permanent damage, risking plant personnel safety and reducing profits. In this context accurate fault diagnosis is fundamental. This work presents an approach for fault diagnosis in industrial process. The diagnostic is performed using classifiers and multivariate data analysis techniques. To improve diagnosis accuracy, faults are clustered by the influence of the variables. Therefore, a single classifier is replaced by multiple local classifiers. Using a single classifier for all faults can make the task of classification more complex and reduce the accuracy of the diagnosis, while local classifiers may be less complex and have a greater power of discrimination among different faults. In addition, to simplify the data to be analyzed by the classifiers, the proposed approach uses feature extraction to analyze the behavior of the process during the occurrence of a fault. The activities are carried out using MATLAB and the approach is applied to three case studies: the classification of time series available in databases in the literature, simulations of the Tennessee Eastman Process plant and simulations of a continuous stirred tank reactor.

Lista de Figuras

Figura 1 – Classificação dos algoritmos de diagnóstico.....	14
Figura 2 – Classificação dos métodos baseados em dados históricos de processo	18
Figura 3 – O problema na análise de variáveis individualmente.....	20
Figura 4 – Análise externa.....	32
Figura 5 – Uma árvore de decisão e as regiões de decisão do espaço de objetos.	38
Figura 6 – Exemplo de <i>random forest</i>	39
Figura 7 – Treinamento dos classificadores	49
Figura 8 – Aplicação dos classificadores.	49
Figura 9 – Comparação com os resultados de Li et. al. (2016).	54
Figura 10 – Comparação a metodologia de maior média de Li et. al. (2016).	54
Figura 11 – Processo Tennessee Eastman	56
Figura 12 – Matriz de confusão para o subconjunto de falhas.	61
Figura 13 – Falhas presentes em cada grupo	63
Figura 14 – Matriz de confusão para o todas as de falhas.	65
Figura 15 – Visualização bidimensional das características utilizando t-SNE para as falhas IDV(3), IDV(9) e IDV(15).....	66
Figura 16 – O processo CSTR.....	68
Figura 17 – Alteração dos <i>setpoints</i>	71
Figura 18 – Resíduos obtidos.	72
Figura 19 – Resíduos obtidos.	72
Figura 20 – Teste dos classificadores utilizando as medições das variáveis em múltiplos pontos de operação.	75
Figura 21 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando as medições das variáveis.	76
Figura 22 – Teste dos classificadores utilizando as contribuições das variáveis em múltiplos pontos de operação.	77
Figura 23 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando as contribuições das variáveis.	78
Figura 24 – Teste dos classificadores utilizando os desvios das variáveis em múltiplos pontos de operação.	79
Figura 25 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando os desvios das variáveis.....	79
Figura 26 – Teste dos classificadores com resíduos em múltiplos pontos de operação.	81
Figura 27 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando os resíduos.	81
Figura 28 – Teste dos classificadores utilizando as contribuições dos resíduos em múltiplos pontos de operação	83
Figura 29 – Teste dos classificadores treinados com dados de falhas de todos os pontos de operação utilizando as contribuições dos resíduos.	83
Figura 30 – Comparação das acurácias obtidas para diferentes sinais e pontos de operação	84

Lista de Tabelas

Tabela 1 – Valores de M	26
Tabela 2 – Bases de dados	51
Tabela 3 – Comparação dos resultados	52
Tabela 4 – Variáveis manipuladas do processo	56
Tabela 5 – Variáveis medidas do processo	57
Tabela 6 – Variáveis medidas no fluxo de alimentação	57
Tabela 7 – Variáveis medidas no gás de purga	57
Tabela 8 – Variáveis medidas no fluxo de produto	58
Tabela 9 – Falhas do processo	58
Tabela 10 – Resultado do agrupamento para um subconjunto de falhas	60
Tabela 11 – Acurácia da metodologia para um subconjunto de falhas utilizando as medições das variáveis	60
Tabela 12 – Acurácia da metodologia para um subconjunto de falhas utilizando as contribuições	60
Tabela 13 – Acurácia da metodologia utilizando as medições das variáveis para todas as falhas	63
Tabela 14 – Acurácia da metodologia utilizando os dados de contribuição para todas as falhas	64
Tabela 15 – Variáveis do simulador	69
Tabela 16 – Lista de falhas do processo	69
Tabela 17 – <i>Setpoints</i> para os pontos de operação escolhidos	70
Tabela 18 – Resultado do agrupamento de falhas para o CSTR	73

Sumário

1	Introdução.....	13
1.1	Justificativa	15
1.2	Objetivos e estrutura da dissertação.....	16
2	Revisão da Literatura	17
2.1	Métodos de detecção de falha baseados em dados históricos.....	17
2.1.1	Análise univariada	18
2.1.2	Análise multivariada.....	20
2.1.2.1	Análise por componentes principais – PCA	21
2.1.2.2	Detecção de falha utilizando PCA.....	24
2.1.2.3	Análise dinâmica de componentes principais – DPCA.....	26
2.1.2.4	Direção de falha.....	27
2.2	Diagnóstico de falhas.....	29
2.2.1	Gráficos de contribuição.....	29
2.3	Análise externa	31
2.4	Extração de características	33
2.4.1	Um ponto	33
2.4.2	Dois pontos consecutivos	34
2.4.3	Três pontos consecutivos.....	35
2.4.4	Subconjuntos	35
2.5	Classificadores	36
2.5.1	Árvores de decisão.....	37
2.5.2	<i>Random forest</i>	39
2.6	Agrupamento de dados	40
2.6.1	K-médias (<i>k-means</i>)	41
2.6.2	Validação do número de grupos	42
2.7	Visualização de dados em alta dimensão.....	43
3	Metodologia proposta.....	45
3.1	Obtenção dos modelos de PCA e DPCA	46
3.2	Detecção de falha.....	46
3.3	Agrupamento de falhas	47
3.4	Dados para os classificadores	47
3.5	Treinamento dos classificadores	48

3.6	Aplicação	49
4	Estudos de caso	50
4.1	Séries temporais da literatura.....	50
4.2	Processo Tennessee Eastman (TEP)	55
4.2.1	Descrição do simulador	55
4.2.2	Deteccção e diagnóstico de falhas.....	59
4.2.3	Teste usando um subconjunto das falhas.....	59
4.2.4	Todas as falhas	62
4.3	Reator tanque agitado contínuo	67
4.3.1	Descrição do simulador	68
4.3.2	Análises realizadas	70
4.3.3	Criação do modelo de variáveis externas	70
4.3.4	Falhas em múltiplos pontos de operação.....	73
4.3.5	Treinamento dos classificadores.....	73
4.3.6	Classificadores treinados com as medições das variáveis	74
4.3.7	Classificadores treinados com as contribuições das variáveis	76
4.3.8	Classificadores treinados com os desvios das variáveis.....	78
4.3.9	Classificadores treinados com os resíduos	80
4.3.10	Classificadores treinados com as contribuições dos resíduos	82
5	Conclusão.....	86
6	Referências.....	87

1 Introdução

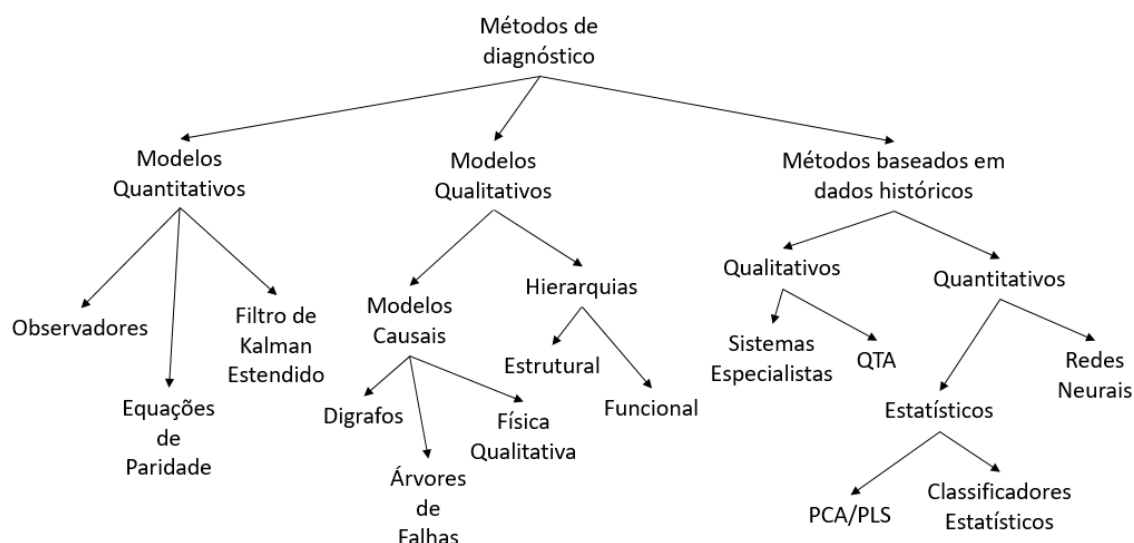
Avanços tecnológicos na indústria vêm resultando em processos cada vez mais complexos, bem como sistemas e produtos com grandes desafios em projeto, análise, fabricação e gerenciamento de suas operações (MAURYA; RENGASWAMY; VENKATASUBRAMANIAN, 2007).

Detecção e diagnóstico de falhas são problemas importantes na engenharia de processos, sendo os componentes centrais da gestão de eventos anormais. A detecção precoce e o diagnóstico de falhas do processo podem ajudar a evitar a progressão da anormalidade, evitando a quebra de equipamentos e paradas indesejadas da planta, reduzindo custos (VENKATASUBRAMANIAN; RENGASWAMY; KA, 2003).

Venkatasubramanian et al. (2003) classificou os algoritmos de detecção e diagnóstico de falha em três grupos, são eles: métodos baseados em modelos qualitativos, modelos quantitativos e em dados.

Nos métodos baseados em modelos quantitativos, conhecimento a priori do processo é utilizado na criação de modelos baseados nas leis físicas que regem o comportamento do processo estudado, na forma de funções matemáticas que correlacionam entradas e saídas do sistema. Em contraste, no modelo qualitativo, o comportamento do processo pode ser analisado através termos de funções qualitativas centradas em diferentes unidades de um processo, através de modelos identificados do processo. Nos métodos baseados em dados históricos de processo, apenas a disponibilidade de grande quantidade de dados é assumida. A Figura 1 ilustra a classificação dos métodos citados para detecção e diagnóstico de falhas.

Figura 1 – Classificação dos algoritmos de diagnóstico



Fonte: Adaptado de Venkatasubramanian et. al (2003).

Métodos de projeção multivariada têm sido utilizados para analisar grandes conjuntos de dados de processos industriais com o objetivo de melhorar o desempenho do processo e qualidade dos produtos. Melhorias operacionais envolvem o desenvolvimento de métodos mais sofisticados para a análise de dados históricos, solução de problemas, monitoramento e otimização de processos e produtos, detecção e diagnóstico de falhas. O uso de modelos estatísticos multivariados é particularmente atraente porque reduz a dimensionalidade dos dados analisados, permitindo uma fácil representação gráfica, interpretação e tratamento de dados faltantes que surgem, por exemplo, de problemas em sensores e falha de comunicação com o *software* de aquisição de dados (KOURTI, 2002).

Dentre as técnicas existentes, a Análise por Componentes Principais (*Principal Component Analysis* – PCA) é muito utilizada devido a facilidade de sua aplicação, elevada difusão na literatura e alta capacidade de compressão de dados, simplificando o processo analisado através da análise de um número reduzido de componentes. Além de permitir o monitoramento estatístico de processos, a metodologia também fornece uma ferramenta para diagnóstico conhecida como gráficos de contribuição, em que são identificadas as variáveis que influenciaram na falha (ALCALA; QIN, 2009).

Após uma falha ser detectada, deve ser realizado o diagnóstico, ou seja, encontrar a causa da falha e sua localização. Gráfico de contribuição é a ferramenta que tem como objetivo identificar quais variáveis influenciam na detecção da falha. Esses gráficos são

baseados na ideia de que variáveis com maiores contribuições para o índice de detecção de falha são provavelmente as variáveis responsáveis pela falha (ALCALA; QIN, 2009).

Quando há abundância de dados históricos de falhas, métodos de classificação e agrupamento podem ser aplicados para auxiliar no diagnóstico (QIN, 2003). A tarefa de diagnóstico consiste em determinar o tipo, tamanho e localização da falha (ISERMANN, 1997). Técnicas de aprendizado de máquina podem ser aplicadas para capturar associações de dados para realizar o diagnóstico de falhas. Algumas técnicas da literatura foram aplicadas para esta tarefa, como os classificadores Bayesianos de rede por Verron et al. (2006), redes neurais artificiais (*Artificial Neural Network* – ANN) por Zhang (2006), máquinas de vetor de suporte (*Support Vector Machine* – SVM) por Widodo e Yang (2007) e análise lógica de dados (*Logical Analysis of Data* – LAD) (RAGAB et al., 2017).

1.1 Justificativa

Em processos petroquímicos estima-se que há uma perda anual de aproximadamente 20 bilhões de dólares devido a falhas. Logo, diagnosticar uma falha enquanto o processo ainda opera em uma região na qual é possível realizar a correção do problema sem que seja necessária uma parada imprevista com perda de produção, ajuda a reduzir gastos e auxilia na operação segura da planta (VENKATASUBRAMANIAN; RENGASWAMY; KA, 2003). Caso não seja possível corrigir a falha antes da necessidade de uma parada imprevista no equipamento, ainda é necessário determinar o que gerou a falha, para que sejam realizadas as correções necessárias no processo ou nos equipamentos e tal falha não volte a ocorrer.

Com a grande quantidade de dados disponíveis em processos automatizados, busca-se a aplicação de uma metodologia que utilize tais dados para determinar a causa da falha. Mesmo com a gama de trabalhos já disponíveis na literatura, um método capaz de tratar desde a detecção de uma falha até o seu diagnóstico, faz-se necessário. Uma metodologia com este objetivo é proposta neste trabalho, utilizando algoritmos já consolidados na literatura para a detecção de falhas, agrupamento e classificação de falhas.

1.2 Objetivos e estrutura da dissertação

Esta dissertação tem por objetivo desenvolver uma metodologia para diagnóstico de falhas que reúne análise estatística multivariada, aprendizagem de máquina e agrupamento de falhas com base em suas similaridades. Para cada grupo de falhas conhecidas, um classificador especializado é treinado para classificar tais falhas. A abordagem usa a extração de características dos dados de processo, após a detecção de uma falha, para treinar os classificadores capazes de discriminar e diagnosticar corretamente as falhas.

A motivação do uso de múltiplos classificadores, se baseia no fato de que um único classificador necessitaria de uma complexidade maior para realizar a separação correta das diferentes falhas em um processo. Uma vez que classificadores locais podem ser mais simples por se especializarem em um subconjunto de falhas, tornando o diagnóstico mais preciso.

Este trabalho é organizado da seguinte forma: no Capítulo 2 é feita uma revisão bibliográfica apresentando as técnicas utilizadas no trabalho. O Capítulo 3 descreve a metodologia proposta para diagnóstico de falhas. No Capítulo 4 a metodologia proposta é aplicada a três estudos de caso. As conclusões finais e sugestões para trabalhos futuros são apresentadas no Capítulo 5.

2 Revisão da Literatura

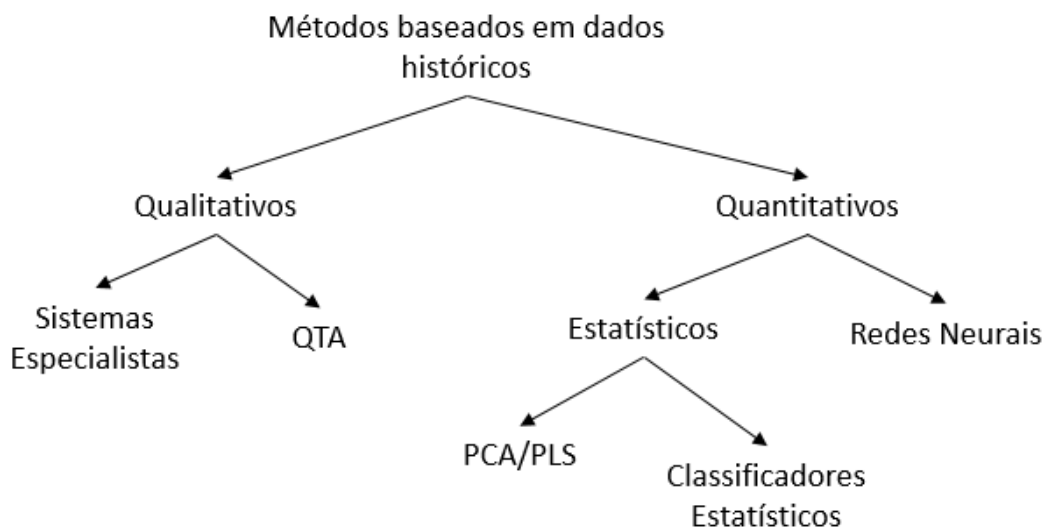
Nesse capítulo são apresentados os conceitos e ferramentas necessárias para o desenvolvimento desse trabalho. A seção 2.1 apresentará os métodos utilizados para detecção de falhas. A seção 2.2 mostrará as metodologias utilizadas para diagnóstico. A seção 2.3 explicará os algoritmos de agrupamento e a suas aplicações. Na seção 2.4 será discutido e apresentado o uso de extração de características para caracterizar séries temporais. A seção 2.5 mostrará os conceitos do classificador aplicado.

2.1 Métodos de detecção de falha baseados em dados históricos

Ao contrário das metodologias baseadas em modelos de processo, no caso de métodos baseados em dados históricos nenhum conhecimento do sistema é necessário, apenas a disposição de uma quantidade grande de dados torna possível detectar situações de funcionamento anormal da planta estudada (VENKATASUBRAMANIAN et al., 2003). Os métodos baseados em dados históricos podem ser divididos em quantitativos e qualitativos.

Dois dos principais métodos qualitativos são o uso de sistemas especialistas e análise de tendência. Métodos quantitativos são de modo geral classificados como métodos não-estatísticos e estatísticos. Redes neurais são uma importante classe de métodos não estatísticos. No caso de métodos estatísticos técnicas como a PCA, Mínimos Quadrados Parciais (*Partial Least Squares* – PLS) e classificação por padrões são as principais formas de detecção de falhas. Na Figura 2 são apresentadas as diferentes maneiras em que conhecimento pode ser extraído dos dados de processo.

Figura 2 – Classificação dos métodos baseados em dados históricos de processo



Fonte: Adaptado de Venkatasubramanian et al. (2003).

Entre os métodos qualitativos, os sistemas especialistas são destinados a solucionar problemas em áreas específicas de conhecimento, buscando um desempenho equiparável ao de especialistas humanos na execução de suas tarefas. Já a análise de tendências (*Qualitative Trend Analysis* – QTA) busca modelar a tendência temporal dos dados e pode ser utilizada para explicar vários eventos importantes que ocorrem no processo, realizar diagnóstico de anormalidades e prever estados futuros.

Os métodos quantitativos essencialmente transformam o problema de diagnosticar uma falha em um problema de reconhecimento de padrões, onde o objetivo é a classificação dos dados em classes pré-determinadas. Abordagens como a PCA extraem informações sobre as principais tendências presentes nos dados e os representam através de um número reduzido de componentes relevantes, as componentes principais. Redes neurais assumem uma forma funcional de regra de decisão, parametrizando assim o classificador.

2.1.1 Análise univariada

Controle de qualidade foi uma das primeiras tentativas de utilizar análise estatística em monitoramento *on-line* e detecção de falhas. As cartas de controle (SHEWART, 1931) foram propostas para realizar a análise univariada de variáveis de qualidade de processo.

A abordagem por cartas de controle é baseada na suposição que um processo submetido a sua variabilidade natural irá permanecer em um estado de controle estatístico no qual as variáveis de processo e/ou de produto irão permanecer próximas aos seus valores desejados.

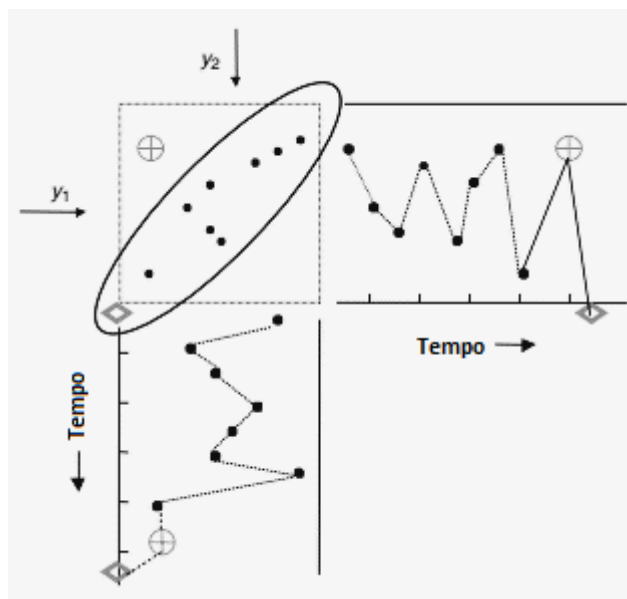
Um processo que opere em condições estáveis também é previsível, ou seja, uma vez que sua variação natural seja determinada é possível construir modelos estatísticos para suas variáveis. O processo é considerado estacionário, a não ser que ações externas sejam tomadas para alterá-lo, como ocorrência de falhas ou mudanças de ponto de operação. Assim se torna possível monitorar as variáveis do processo com base nos seus padrões já estabelecidos em funcionamento normal.

Shewart desenvolveu a metodologia de cartas de controle para que fosse possível distinguir entre alterações inerentes do processo e casos especiais de variação da variável analisada. Também buscou fornecer uma regra para minimizar o risco de reação a eventos em que a alteração dos dados é de fato uma situação normal. Para isso, após análises empíricas ele determinou os limites de controle superior (LCS) e inferior (LCI), a partir de $\pm 3\sigma$ da média das amostras, em que σ é o desvio padrão da variável analisada em produção regular. Em casos onde a distribuição dos dados é normal esses limiares correspondem a um intervalo de confiança de 99%. A violação de um desses limites representa uma mudança inesperada no funcionamento do processo, indicando a ocorrência de uma falha.

A crescente demanda pela melhora na qualidade dos produtos e confiabilidade do processo tem levado ao uso extensivo de cartas de controle estatístico de processo (CEP). Embora o conceito por trás das cartas de controle ainda seja válido, esse método não permite analisar processos complexos. Em casos em que as variáveis monitoradas não são independentes, o uso de cartas de controle univariadas pode levar a conclusões equivocadas, devido ao fato de métodos univariados não considerarem a correlação entre as variáveis.

A análise comparativa entre métodos uni e multivariáveis, representada Figura 3, é utilizada na literatura para ilustrar a deficiência do uso de duas cartas de controle para variáveis correlacionadas. Nela são analisadas duas variáveis (y_1 e y_2) mostradas de forma conjunta (canto superior esquerdo) e as mesmas amostras são exibidas individualmente com seus limiares de controle.

Figura 3 – O problema na análise de variáveis individualmente



Fonte: Adaptado de Kourti (2002).

Enquanto o processo opera em condições normais y_1 e y_2 seguem uma distribuição multivariada normal. Os pontos no gráfico conjunto mostram que as observações de y_1 e y_2 são correlacionadas. A elipse representa um intervalo de confiança de $(1 - \alpha)\%$, sendo α o limite de confiança da distribuição, ou seja, mesmo com o sistema operando em condições normais uma fração α de amostras estará fora da elipse.

O ponto indicado por \oplus está fora do intervalo de confiança conjunto, porém quando são analisadas as cartas de controle de cada uma das variáveis não há indicação que esse ponto represente uma condição anormal do sistema. O ponto indicado por \diamond será rejeitado pelas análises univariadas enquanto na realidade é um ponto de funcionamento normal quando visto na análise multivariada. Esse problema se torna mais complexo com o aumento no número de variáveis de um processo e tornando necessária a análise multivariada.

2.1.2 Análise multivariada

Examinar variáveis separadamente torna a interpretação dos sinais mais difícil, podendo resultar em conclusões equivocadas a respeito do funcionamento do sistema. Abordagens multivariadas conseguem tratar os dados simultaneamente e extrair

informações corretas das variações do processo, considerando como as variáveis interagem entre si.

Técnicas estatísticas multivariadas são ferramentas poderosas capazes de comprimir os dados e reduzir a dimensionalidade dos mesmos, de tal forma que apenas as informações essenciais são mantidas, tornando a análise do processo mais simples.

2.1.2.1 Análise por componentes principais – PCA

A PCA foi proposta em Pearson (1901) e posteriormente desenvolvida por Hotelling (1947) com o objetivo de definir um grupo de componentes principais que consistem na combinação linear dos dados originais, de forma que a primeira componente principal represente a maior variação dos dados, a segunda componente principal a segunda maior variação e assim em diante. As componentes são ortogonais entre si e preservam a correlação existente entre as variáveis do processo (ALDRICH; AURET, 2016).

O conjunto completo de componentes principais contém a variância total dos dados analisados. A maior parte da variância pode ser mantida analisando um grupo menor de componentes, reduzindo a dimensionalidade dos dados. Utilizando dados históricos de funcionamento normal do sistema é possível construir um modelo que leva em consideração as relações entre as variáveis (CHEN, 2005).

Dado um conjunto de dados de treinamento, $X = [x_1 \ x_2 \ \dots \ x_m]^T \in \mathbb{R}^{n \times m}$ com m variáveis e n amostras, separados com o sistema em condições normais de operação. A matriz de covariância amostral S dos dados é definida por

$$S = \frac{1}{n-1} X^T X = P \Lambda P^T \quad (1)$$

e a PCA calcula um conjunto de vetores ortogonais, chamados vetores de carregamento, a partir da decomposição em valores singulares de S , em que $P \in \mathbb{R}^{m \times m}$ é uma matriz unitária, e suas colunas são os vetores de carregamento. A matriz diagonal $\Lambda \in \mathbb{R}^{m \times m}$ contém os autovalores reais não-negativos ordenados em ordem decrescente ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$) onde cada autovalor representa a variância capturada pelo vetor de carregamento correspondente (CHIANG; RUSSEL; BRAATZ, 2000).

A correta escolha dos dados de treinamento é uma importante etapa do método, uma vez que eles reproduzem a variabilidade natural dos dados. Para evitar que tendências e

diferenças de escalas das variáveis interfiram no cálculo das componentes principais é importante realizar o tratamento prévio dos dados, como a retirada de médias temporais e a normalização pelo desvio padrão.

O conceito de redução de dimensionalidade utilizando a PCA foca na seleção ótima do número de componentes principais (l), de forma que são selecionadas as colunas de matriz de carregamento correspondentes aos vetores associados aos primeiros l autovalores, separando a matriz de carregamento da seguinte forma:

$$P = [\hat{P} \ \tilde{P}] \quad (2)$$

onde $\hat{P} \in \mathbb{R}^{m \times l}$ é a representação da matriz de carregamento no subespaço das componentes principais e $\tilde{P} \in \mathbb{R}^{m \times (m-l)}$ é a representação da matriz de carregamento no subespaço residual, a mesma decomposição pode ser realizada para a matriz Λ .

A PCA projeta, a partir de transformações lineares, os dados das observações de X em uma nova matriz de dimensão reduzida, $T \in \mathbb{R}^{n \times l}$ denominada matriz de pontuações e a projeção de T no espaço m -dimensional é a matriz de valores previstos \hat{X} :

$$T = X\hat{P} \quad (3)$$

$$\hat{X} = T\hat{P}^T \quad (4)$$

A diferença entre X e \hat{X} é chamada de matriz residual (E) e captura as variações no espaço de observação abrangido pelos vetores de carregamento associados aos $m - l$ menores autovalores.

$$E = X - \hat{X} = \tilde{P}\tilde{P}^T X \quad (5)$$

O subespaço representado pela matriz residual contém uma baixa relação sinal-ruído, e a remoção desse subespaço pode produzir uma representação mais precisa \hat{X} . A escolha do número de componentes principais é um passo de extrema importância para a aplicação da PCA e deve ser realizada de forma adequada. Em casos em que poucas componentes principais são escolhidas, o modelo obtido não representará a variabilidade do processo. Contudo, quando um número grande de componentes é selecionado, as componentes com menores valores singulares serão incluídas, o que terá um enorme efeito no cálculo da estatística T^2 , pois os quadrados dos valores singulares são invertidos na Equação (7).

As técnicas mais utilizadas para o cálculo do número de componentes principais são (CHIANG; RUSSEL; BRAATZ, 2000):

- Variância percentual acumulada: esse método determina o número de componentes principais através do cálculo do menor número de vetores de carregamento necessários para explicar um percentual escolhido da variância total do sistema;
- *Scree test*: nesse teste se assume que a variância referente ao ruído tem um perfil linear. O número de componentes selecionadas é determinado localizando o perfil da variância que não é mais linear. A identificação desse valor pode ser ambígua, dificultando assim sua automatização;
- Análise paralela: determina a dimensionalidade final através da comparação do perfil da variância (matriz Λ) com o perfil obtido considerando que as variáveis sejam independentes. A redução da ordem é determinada como o ponto de cruzamento entre os dois perfis de variância;
- Validação cruzada: é realizada utilizando o valor previsto para soma dos quadrados (*prediction residual sum of squares* – PRESS)

$$PRESS(i) = \frac{1}{mn} \|X - \hat{X}\|_F^2 \quad (6)$$

onde i representa o número de componentes retidas utilizado no cálculo de \hat{X} e $\|\cdot\|_F$ é a norma de Frobenius. Para a aplicação dessa técnica o conjunto de treinamento é dividido em grupos. A estatística PRESS é calculada para várias dimensões no espaço de pontuações (i) em um grupo. Isso deve ser repetido para cada grupo, e o valor de i associado com a menor média da estatística PRESS determina a o número de componentes selecionadas.

Abdi e Williams (2010) listam o objetivo do uso da PCA como:

- i. Extrair as informações mais importantes de um conjunto de dados;
- ii. Comprimir o conjunto de dados mantendo apenas as informações mais relevantes;
- iii. Simplificar a descrição do conjunto de dados;

- iv. Analisar a estrutura das medições e das variáveis.

2.1.2.2 Detecção de falha utilizando PCA

Detecção de falhas é o primeiro passo no monitoramento multivariado de processo (QIN, 2003). A utilização da PCA permite o acompanhamento do processo através do cálculo de estatísticas multivariadas. Entre essas estatísticas se destacam a T^2 de Hotelling e o erro quadrático de predição (*square prediction error* – SPE), também conhecido como estatística Q . Essas estatísticas projetam a variabilidade do processo nos subespaços de componentes principais e residuais, respectivamente. Devido à natureza complementar dessas estatísticas, um índice combinado (ϕ) pode ser utilizado para observar simultaneamente os dois subespaços resultantes da aplicação da PCA (YUE; QIN, 2001).

A variação do processo no subespaço de componentes principais para uma determinada amostra normalizada de dados, x , pode ser monitorada através do índice T^2 definido por

$$T^2 = x^T \hat{P} \hat{\Lambda}^{-1} \hat{P}^T x = x^T D x \quad (7)$$

em que a matriz $\hat{\Lambda}$ contém os valores singulares referentes as componentes principais selecionadas.

Em condições normais de funcionamento do processo, quando a matriz de covariância é estimada a partir da covariância amostral dos dados, o índice T^2 está relacionado a distribuição estatística F (CHIANG; RUSSEL; BRAATZ, 2000). Essa distribuição pode ser utilizada para calcular o limiar de controle de T^2 para um dado nível de significância α através de

$$\tau_{\alpha}^2 = \frac{l(n-1)}{n-l} F_{l,n-1;\alpha} \quad (8)$$

sendo, $F_{l,n-1;\alpha}$ representa a distribuição F com l e $n-1$ graus de liberdade para um nível de significância α . Com isso considera-se que o sistema está funcionando de forma adequada quando a estatística calculada está abaixo do limiar de controle.

A estatística Q mede a projeção de um vetor normalizado de amostras x no subespaço residual sendo calculada a partir da Equação (9), e o processo é considerado em funcionamento normal quando é satisfeita a desigualdade dada pela Equação (10).

$$Q = \|\tilde{x}\|^2 = x^T \tilde{P} \tilde{P}^T x = x^T \tilde{C} x \quad (9)$$

$$Q \leq \delta_\alpha^2 \quad (10)$$

em que δ_α^2 denota o limiar de controle para a estatística com um nível de significância α .

Jackson e Howard (1979) desenvolveram uma expressão para o cálculo de δ_α^2 :

$$\delta_\alpha^2 = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (11)$$

sendo c_α o limiar de uma distribuição normal padrão com intervalo de confiança $(1 - \alpha)\%$ e os outros elementos são calculados por:

$$\theta_i = \sum_{j=l+1}^m \lambda_j^i \quad (12)$$

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \quad (13)$$

Devido à natureza complementar das duas estatísticas, Yue e Qin (2001) propõem o uso de um índice combinado, φ , para simplificar a tarefa de detecção de falha. Esse índice incorpora de uma forma balanceada ambas as estatísticas

$$\varphi = \frac{Q}{\delta_\alpha^2} + \frac{T^2}{\tau_\alpha^2} = x^T \Phi x \quad (14)$$

$$\Phi = \frac{\tilde{C}}{\delta_\alpha^2} + \frac{D}{\tau_\alpha^2} \quad (15)$$

O limiar de controle para o índice combinado segue aproximadamente uma distribuição estatística qui-quadrado de forma que

$$\zeta^2 = g^\varphi \chi_\alpha^2(h^\varphi) \quad (16)$$

$$g^\varphi = \left(\frac{l}{\tau_\alpha^4} + \frac{\theta_2}{\delta_\alpha^4} \right) / \left(\frac{l}{\tau_\alpha^2} + \frac{\theta_1}{\delta_\alpha^2} \right) \quad (17)$$

$$h^\varphi = \left(\frac{l}{\tau_\alpha^2} + \frac{\theta_1}{\delta_\alpha^2} \right)^2 / \left(\frac{l}{\tau_\alpha^4} + \frac{\theta_2}{\delta_\alpha^4} \right) \quad (18)$$

É possível notar uma semelhança entre as Equações (7), (9) e (14), possibilitando a generalização das equações dos diferentes índices no formato:

$$\text{Índice} = x^T M x \quad (19)$$

em que M é a matriz característica do índice que se deseja analisar. Seus valores são apresentados na Tabela 1.

Tabela 1 – Valores de M

<i>Índice</i>	<i>M</i>
Q	\tilde{C}
T^2	D
φ	Φ

Fonte: Adaptado de Joe e Qin (2003)

O monitoramento do processo é realizado acompanhando um ou mais índices. Uma falha é detectada quando um dos limiares de controle é violado. Nesse trabalho será utilizado o índice combinado para detecção de ocorrência de falhas.

2.1.2.3 Análise dinâmica de componentes principais – DPCA

Técnicas de monitoramento baseadas na PCA assumem implicitamente que as medições realizadas em um determinado instante são independentes das medições realizadas em instantes anteriores. Tipicamente em processos industriais essa suposição só é verdadeira para amostras com longos períodos de amostragem, por exemplo, 2 a 12 horas (CHIANG; RUSSEL; BRAATZ, 2000).

Quando são considerados os efeitos dinâmicos e controle em malha fechada em processos industriais, as medições das variáveis em diferentes instantes não são independentes, o que indica que essas medições são correlacionadas temporalmente. Para capturar as relações dinâmicas entre as variáveis, Ku, Storer e Georgakis (1995) propõem uma variação da PCA chamada de Análise Dinâmica de Componentes Principais (*dynamic principal component analysis* – DPCA) para que sejam consideradas as correlações temporais entre as variáveis. Isso é realizado levando em consideração não apenas os dados medidos em um determinado instante, mas também as d medições anteriores, o que resulta na matriz aumentada

$$Z = \begin{bmatrix} x_d & x_{d-1} & \dots & x_1 \\ x_{d+1} & x_d & \dots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{n+d-1} & x_{n+d-2} & \dots & x_n \end{bmatrix} \quad (20)$$

em que $x_k \in \mathbb{R}^m$ é um vetor com a medida de todas as variáveis em um dado instante, d é o número de atrasos aplicados a matriz de dados, sendo $Z \in \mathbb{R}^{n \times md}$. Aplicando PCA à matriz aumentada, um modelo multivariado autorregressivo (AR) é extraído dos dados. Li et.al. (2014) demonstra a equivalência entre PCA e DPCA para monitoramento de processos. A utilização do DPCA se resume em aplicar as metodologias descritas na seção 2.1.2.1 a matriz aumentada Z ao invés da matriz de dados X .

Nesse trabalho o número de atrasos analisados é calculado através do valor singular chave (*key singular value* – KSV) (RATO; REIS, 2013). Esse método assume a existência de um total de m relações dinâmicas entre as variáveis a serem identificadas, sendo m o número total de variáveis, e não se sabe de antemão a ordem das relações temporais entre as medições. Os passos do método proposto são:

1. Inicia-se com $d = 0$;
2. É criada a matriz aumentada Z ;
3. É realizada a decomposição da matriz de covariância de Z : $S_Z = UAV^T$;
4. $KSV(d) = a_{sd+1}$, corresponde ao elemento $sd + 1$ da matriz A ;
5. Se $d > 0$, então $KSVR(d) = KSV(d)/KSV(d - 1)$;
6. Se $d < d_{max}$ (d_{max} é um valor escolhido máximo de atrasos), $d = d + 1$ e retorna-se ao passo 2, caso contrário;
7. Normaliza-se KSV e $KSVR$;
8. Determina-se d como $\arg \min_{d \in [1; d_{max}]} \sqrt{KSV_N(d)^2 KSVR_N(d)^2}$.

2.1.2.4 Direção de falha

Qin (2003) define a direção da falha como vetores ou matrizes ortonormais, para falhas em uma única variável e para falhas em múltiplas variáveis, respectivamente. A direção de falha indica quais variáveis estão em falha.

Com o intuito de realizar o agrupamento de falhas que influenciaram variáveis semelhantes, nesse trabalho o conceito da direção de falha é utilizado. Para realizar o agrupamento de falhas, a direção de falha será representada por um vetor binário (ξ), que

indicará quais variáveis foram influenciadas após a ocorrência de uma determinada falha. Possibilitando o agrupamento de falhas a partir das variáveis impactadas.

Em casos em que os históricos de falhas estão disponíveis, a direção de falha representa o subespaço de variáveis onde o processo desvia de seu funcionamento normal (QIN, 2003).

Liu e Chen (2014) apresentam um algoritmo para isolar variáveis em falha através da análise do índice combinado. O algoritmo remove variáveis da análise até que, ao realizar uma análise multivariada nas variáveis restantes, a falha não seja mais detectada pelo índice combinado. Dessa forma, dois grupos de variáveis podem ser separados. O primeiro grupo compreende as variáveis que não tiveram qualquer influência na falha. O outro grupo é composto pelas variáveis que influenciaram na falha, portanto, se o índice combinado for calculado usando essas variáveis, a falha seria detectada. A abordagem segue os passos:

1. A partir do modelo de PCA é realizado o monitoramento do processo utilizando a estatística combinada;
2. Após detectada uma falha são calculadas as contribuições das variáveis, como será demonstrado na seção 2.2.1;
3. A variável com maior contribuição é adicionada ao grupo de variáveis que possuíram influência na falha e em seguida é removida da análise;
4. Atualiza-se o modelo de PCA, removendo as variáveis que influenciaram na falha, e retorna ao passo 1 até que a análise realizada com as variáveis resultantes não detecte mais a falha;
5. Separam-se as variáveis nos dois grupos citados anteriormente.

Após a separação dos grupos é criado o vetor ξ de modo que as variáveis no grupo um (que não influenciaram na falha) são representadas pelo número zero e as variáveis no grupo dois (que influenciam na falha) são representadas pelo número um. Se os dados de várias falhas estiverem disponíveis para análise, as direções de falhas podem ser calculadas e usadas para isolar falhas que influenciam um conjunto de variáveis semelhantes. Consequentemente, isso permite o uso de algoritmos de agrupamento para agrupar falhas, simplificando assim a tarefa de diagnóstico.

2.2 Diagnóstico de falhas

Uma vez que uma falha tenha sido detectada é importante determinar qual variável ou variáveis a causaram. Diagnosticar uma falha pode ser uma tarefa extremamente complexa quando o processo possui muitas variáveis ou é altamente integrado. Algumas das variáveis podem se afastar da sua referência por um curto período de tempo, e devido as malhas de controle do processo retornarem ao normal mesmo que a falha persista. Esse tipo de comportamento pode ocultar a falha, tornando extremamente difícil a automatização de um algoritmo para diagnóstico (CHIANG; RUSSEL; BRAATZ, 2000).

2.2.1 Gráficos de contribuição

O gráfico de contribuição é uma ferramenta comumente utilizada para auxílio no diagnóstico de falhas devido a sua simples implementação. Ela se baseia no princípio de que as variáveis com as maiores contribuições são as causadoras da falha (MACGREGOR et al., 1994).

Essa abordagem não requer nenhum conhecimento prévio a respeito do processo analisado, apenas é necessário um modelo de PCA como o utilizado para detecção de falha. A contribuição representa os efeitos das falhas sobre os vetores de dados coletados do processo. Se uma grande quantidade de dados históricos com falhas diferentes estiver disponível, o uso de algoritmos de agrupamento e classificação se torna aplicável sobre os dados de contribuição para diagnóstico automático das falhas (QIN, 2003).

Alcala e Qin (2010) apresentam uma forma de cálculo das contribuições da i -ésima variável para a falha, nos f instantes analisados, na forma:

$$c_i^{indice} = \left(\xi_i^T M^{\frac{1}{2}} x_q \right)^2 \quad (21)$$

sendo ξ_i a i -ésima coluna de uma matriz identidade de tamanho m representando cada variável analisada, x_q é o vetor de variáveis amostradas normalizado com as médias e desvios padrões calculados a partir dos dados de treinamento e $q = t, t + 1, \dots, t + f$. O cálculo da contribuição das m variáveis resulta na matriz de contribuições $C^j \in R^{f \times m}$, onde cada coluna contém a contribuição da i -ésima variável durante os f instantes de falha.

A função desses gráficos nem sempre é indicar as variáveis com maior contribuição, mas sim o comportamento delas, pois falhas diferentes podem ter influências similares em uma mesma variável. Em alguns casos, processos com controle em malha fechada podem levar a diagnósticos equivocados quando analisados apenas os gráficos de contribuição, pois os controladores do processo alteram as variáveis manipuladas para corrigir problemas, mascarando algumas vezes a real responsável pela falha (QIN, 2003).

Embora os gráficos de contribuição tenham uma grande vantagem em relação a sua simples aplicação e entendimento, Westerhius et al. (2000) demonstram que, variáveis em falha podem aumentar a contribuição de variáveis que não possuíram influência para a falha. Como consequência, o efeito da falha é propagado entre as contribuições das variáveis que contribuíram e não contribuíram para a falha, o que pode levar a diagnósticos errados, a esse efeito se dá o nome de efeito *smearing* (QIN, 2003).

Kerkhof et al. (2013) demonstra que o *smearing* está relacionado com a aplicação de técnicas como PCA. A redução da dimensionalidade dos dados e a subsequente expansão para o número original de variáveis, realizada para o cálculo das contribuições, fazem com que as variáveis que estão em falha interajam com as variáveis que não estão, contaminando assim a análise.

Para processos onde a dinâmica do sistema deve ser considerada, Li e Qin (2014) propõem o uso da contribuição utilizando o índice combinado φ , calculado a partir da matriz aumentada (Z) obtida quando se utiliza a técnica de DPCA. As contribuições são calculadas de forma semelhante a contribuição utilizando PCA, porém é utilizada a matriz característica obtida com a DPCA e o vetor de dados z_q , de forma que

$$c_i^{DPCA} = \left(\xi_i^T \Phi^{\frac{1}{2}} z_q \right)^2 \quad (22)$$

O cálculo das contribuições utilizando a equação (22) resulta em uma matriz com as variáveis defasadas no tempo, tornando necessária a redução da dimensão dessa matriz. Isso é realizado somando as contribuições de cada variável as contribuições de suas respectivas variáveis defasadas.

Para evitar erros no diagnóstico de falhas devido à análise apenas dados de contribuição, é desenvolvido o uso desses dados em conjunto com classificadores treinados com falhas conhecidas. Desta forma, não somente a variável com maior

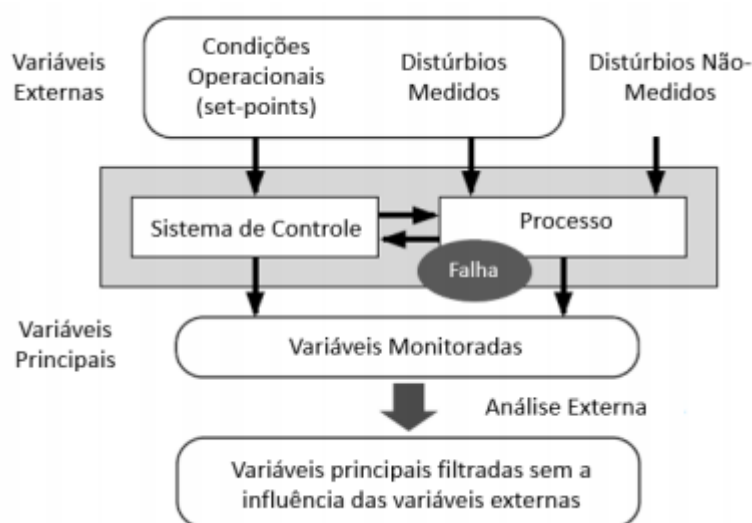
contribuição é considerada, mas sim o comportamento das variáveis para uma determinada falha e, como todas as análises estão baseadas no mesmo modelo de PCA o efeito de *smearing* é inteirado no classificador treinado.

2.3 Análise externa

Processos industriais possuem mais de uma região de operação em que se encontram em estado estacionário. O uso de um modelo multivariado como PCA, treinado para uma única região de operação se torna limitado pois, uma mudança nas condições de operação normal do processo não seria considerada no modelo, indicando a mudança de operação como uma falha.

Kano et al. (2004) propõe o uso de modelos estimados com base nos dados de normalidade em diversas regiões de operação, e chama essa técnica de análise externa. Os autores também afirmam que as condições operacionais de um processo são definidas por fatores externos ao mesmo, como vazões de entrada e *setpoint* de controladores. A técnica então consiste em separar as variáveis do processo em dois grupos: variáveis externas, que comandam o ponto de operação, e variáveis principais, que são afetadas pelas variáveis externas. As variáveis externas são utilizadas para remover a influência da região de operação sobre as variáveis principais através de uma regressão, tornando o monitoramento do processo robusto a mudança de região de operação. Em seguida, é possível calcular os resíduos da regressão possibilitando a aplicação de análises multivariadas, como PCA, para a detecção de falhas. O conceito de análise externa é ilustrado na Figura 4.

Figura 4 – Análise externa.



Fonte: Adaptado de Kano et al. (2004)

Uma matriz de dados em funcionamento normal em diversos pontos de operação $A \in \mathbb{R}^{n \times m}$ pode ser dividida em duas matrizes. A primeira composta apenas pelas medições das variáveis principais (Y) e a segunda composta apenas pelas medições das variáveis externas (U). Como as variáveis principais estão diretamente relacionadas as variáveis externas, é possível decompor a matriz Y da seguinte forma:

$$Y = f(U) + e \quad (23)$$

Em que $f(U)$ representa uma função que relaciona as variáveis externas e principais e e representa os resíduos obtidos da aproximação, ou seja, a parte das variáveis principais que não é explicada pelas variáveis externas. Vários tipos de regressão podem ser utilizados como $f(U)$, os mais comuns são: regressão polinomial e técnicas de aprendizado de máquinas (redes neurais, árvores de decisão e máquinas de vetores de suporte).

Uma vez eliminada a influência das variáveis externas, e calculados os resíduos, se torna possível aplicar técnicas multivariadas como PCA para detecção e diagnóstico de falhas, independente da região de operação em que se encontra o processo. A ocorrência de uma falha se manifesta como uma alteração em alguma das variáveis principais que não é prevista pelo modelo de análise externa. Essa alteração aumentará os resíduos calculados, gerando a violação da estatística analisada. Para diagnóstico podem ser

utilizados diretamente os resíduos como os dados de entrada para os classificadores ou as contribuições dos resíduos para a falha.

2.4 Extração de características

Um grande problema ao se analisar séries temporais é a alta dimensionalidade dos dados. Uma grande quantidade de amostras de dados pode dificultar o processamento de algoritmos (AGRAWAL; FALOUTSOS; SWAMI, 1993). Numa série temporal, não é suficiente analisar cada ponto no tempo sequencialmente, em vez disso, é preciso utilizar janelas deslizantes para caracterizar completamente os sinais (MÖRCHEN, 2003).

Valores consecutivos em séries temporais usualmente são altamente correlacionados. Extração de características é aplicada para comprimir os dados da série temporal, mantendo apenas as informações mais relevantes, removendo pequenas variações e correlações (MÖRCHEN, 2003). As características extraídas podem ser utilizadas em algoritmos de identificação de padrões baseados em similaridades, agrupamento de séries temporais, classificação de séries temporais ou para a detecção de anomalias em séries temporais (DANG; WILKINSON, 2013).

Dang e Wilkinson (2013) propõem uma metodologia para extração de características diretamente dos dados, denominada TimeExplorer. Essa metodologia inclui algumas características estatísticas clássicas, como média, desvio padrão e diferenças, além de levar em consideração também as relações entre dois ou mais pontos de uma série temporal.

A seguir são apresentadas as características extraídas utilizando a metodologia TimeExplorer para uma série de dados $V = v_1, \dots, v_n$ com n observações. As características são classificadas de acordo com o número de observações consecutivas que são consideradas, sendo esse número um, dois, três ou o uso de subconjunto de dados, para extração.

2.4.1 Um ponto

Essas medidas são computadas com base nos dados, onde v_i é a observação no instante i .

1. Média: representa o valor médio da série temporal.

$$Média = \frac{\sum_{i=1}^n v_i}{n} \quad (24)$$

2. Desvio padrão (DP): representa quanto os dados variam em torno da média.

$$DP = \sqrt{\frac{\sum_{i=1}^n (v_i - Média)^2}{n}} \quad (25)$$

2.4.2 Dois pontos consecutivos

Essas características são calculadas baseadas na diferença entre duas observações consecutivas de forma que $dif_i = v_i - v_{i-1}$. As diferenças são divididas em duas classes: aumentos, quando se têm diferenças positivas, e reduções, quando as diferenças calculadas são negativas.

3. Média dos aumentos: é a média dos aumentos encontrados, onde n_a é o número de aumentos.

$$Média_{aumentos} = \frac{\sum_{i=1}^{n_a} aumentos_i}{n_a} \quad (26)$$

4. Média das reduções: é a média das reduções encontradas, onde n_r é o número de reduções.

$$Média_{reduções} = \frac{\sum_{i=1}^{n_r} reduções_i}{n_r} \quad (27)$$

5. Aumento máximo: aumento máximo na série temporal.

$$max_{aumento} = \max(aumentos_{i=1, \dots, n_a}) \quad (28)$$

6. Redução máxima: redução máxima na série temporal.

$$max_{redução} = \max(reduções_{i=1, \dots, n_r}) \quad (29)$$

7. Desvio padrão das diferenças (DP_{dif}): mostra quanta variação existe em torno da média das diferenças.

$$DP_{dif} = \sqrt{\frac{\sum_{i=1}^n (dif_i - Média_{dif})^2}{n-1}} \quad (30)$$

em que

$$Média_{dif} = \frac{\sum_{i=2}^n dif_i}{n-1} \quad (31)$$

2.4.3 Três pontos consecutivos

Essas características são calculadas considerando três observações seguidas de dados. Especificamente são consideradas duas configurações: montanha, quando $v_i > v_{i-1}$ e $v_i > v_{i+1}$ ou vale, quando $v_i < v_{i-1}$ e $v_i < v_{i+1}$. Em outras palavras, uma montanha acontece quando se tem um aumento seguido de uma redução e de maneira oposta um vale acontece quando se tem uma redução seguida de um aumento. Matematicamente tem-se:

$$Montanha_{n_m} = |dif_{i-1}| + |dif_{i+1}| \quad (32)$$

$$Vale_{n_v} = |dif_{i-1}| + |dif_{i+1}| \quad (33)$$

8. Montanha máxima: a montanha máxima em toda a série de dados:

$$max_{montanha} = \max(Montanhas_{i=1..n_m}) \quad (34)$$

9. Vale máximo: o vale máximo da série de dados:

$$max_{vale} = \max(Vale_{i=1..n_v}) \quad (35)$$

2.4.4 Subconjuntos

Seja s o comprimento de um subconjunto em V , uma repartição de observações subsequentes de dados. As características a seguir medem quão bem dois subconjuntos de dados se combinam.

10. Repetibilidade: é a soma das diferenças de uma subsérie comparada com a anterior.

$$Repet = \sum_{i=s+1}^n |v_i - v_{i-s}| \quad (36)$$

11. Periodicidade: Analisa não apenas os padrões de repetição entre os subconjuntos, mas também a variação em cada subsérie.

$$Per = Repet * DP_{dif} \quad (37)$$

Neste trabalho, a extração de características é utilizada para auxiliar o diagnóstico de falhas através da identificação de padrões após a ocorrência de uma falha, seja utilizando os dados do processo ou a contribuição das variáveis para treinar classificadores.

2.5 Classificadores

Um classificador é um algoritmo de aprendizado de máquinas preditivo em que, dado um conjunto de exemplos com classes conhecidas e rotuladas, é construído um estimador. O estimador por sua vez, quando dado um exemplo não rotulado, atribui esse exemplo a uma das possíveis classes (DIETTERICH, 1998).

O diagnóstico de falhas pode ser considerado como um problema de análise ao longo do tempo, das estimativas do classificador, usando o conhecimento sobre as propriedades dos modos de falha do sistema (RENGASWAMY; VENKATASUBRAMANIAN, 2000).

Faceli et al. (2011) dividem os principais métodos de classificação em quatro grupos, são eles:

- Baseados em distâncias: consideram a proximidade entre os dados na realização da classificação. A hipótese básica é que dados similares tendem a estar concentrados em uma mesma região no espaço analisado, enquanto dados não similares estarão em regiões diferentes. Dentre esses métodos se destacam os que utilizam o conceito de vizinhos mais próximos (*nearest neighbors*) como o k-NN;
- Probabilísticos/bayesianos: assumem que a probabilidade de uma amostra pertencer a uma certa classe A , dado um evento B não depende somente da relação entre A e B , mas também da probabilidade de se observar A independentemente de se observar B ;
- Baseados em procura: classificação pode ser formulada como um problema de procura num espaço de soluções possíveis. Esses classificadores buscam a

partir dos dados realizar escolhas no espaço de possíveis classes existentes. Entre eles se destacam as árvores de decisão, florestas aleatórias e modelos baseados em regras;

- Baseados em otimização: buscam pela hipótese que descreve os dados recorrendo à otimização de alguma função. Nesse caso a classificação se torna um problema de minimizar, ou maximizar, uma função objetivo. Desses métodos se destacam as redes neurais artificiais e máquinas de vetores de suporte (*support vector machines* – SVM).

Diversas metodologias de classificação são aplicadas a técnicas de diagnóstico de falhas, baseadas em dados históricos de processo, para classificar dados das séries temporais resultantes após a ocorrência de uma falha (KULKARNI; JAYARAMAN; KULKARNI, 2005).

Trabalhos como Dou e Zhou (2016), Kulkarni et. al. (2005), Lau et. al. 2012 e Eslamloueyan (2010) aplicam, com relativo sucesso, o uso de classificadores como redes neurais, SVM e lógica nebulosa no diagnóstico de falhas em pequenas plantas e simuladores de processos industriais. Os classificadores são treinados com dados de falhas rotulados anteriormente e em seguida são aplicados a uma nova ocorrência de falha.

Nesse trabalho, são utilizadas características do processo, e após a detecção de uma falha é utilizado para classificá-la um classificador do tipo *random forest*. Esse é um método de aprendizagem de máquina que faz uso de diversas árvores de decisão para prever o resultado ou classificar uma série de dados.

2.5.1 Árvores de decisão

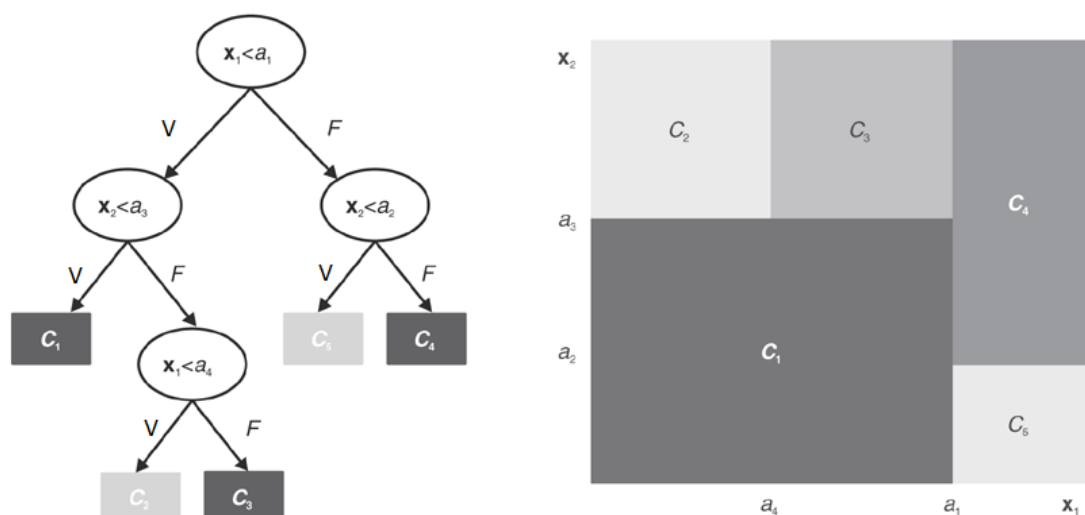
Uma árvore de decisão é um método não paramétrico de aprendizagem de máquinas que usa a estratégia dividir para conquistar. Com ela, um problema complexo é dividido em problemas mais simples, aos quais é recursivamente aplicada a mesma estratégia (FACELI et. al. 2011).

Uma árvore de decisão é um grafo acíclico direcionado em que cada nó pode ser um nó de divisão, com dois ou mais sucessores, ou um nó folha. Faceli et. al. 2011 define esses conceitos como:

- Um nó de divisão contém um teste condicional baseado nos valores dos dados informados ao algoritmo;
- Um nó folha é o rótulo dado pelo algoritmo ao conjunto de dados analisados.

A Figura 5 ilustra um exemplo de árvore de decisão e sua divisão correspondente no espaço definido pelos atributos x_1 e x_2 . A esquerda está a ilustrada a árvore resultante onde as saídas dos nós de divisão são os resultados dos testes condicionais interiores de cada nó. Situações onde a resposta lógica do teste foi falsa são indicadas por F e V indica que a resposta foi verdadeira. Ao final dos testes é realizada a classificação dos dados e a direita da figura está indicado o espaço de decisões da árvore, onde cada classe é indicada pela letra C.

Figura 5 – Uma árvore de decisão e as regiões de decisão do espaço de objetos.



Fonte: Faceli et al. 2011

As árvores de decisão possuem diversas vantagens quando comparadas a outros métodos de classificação (FACELI et al. 2011), são elas:

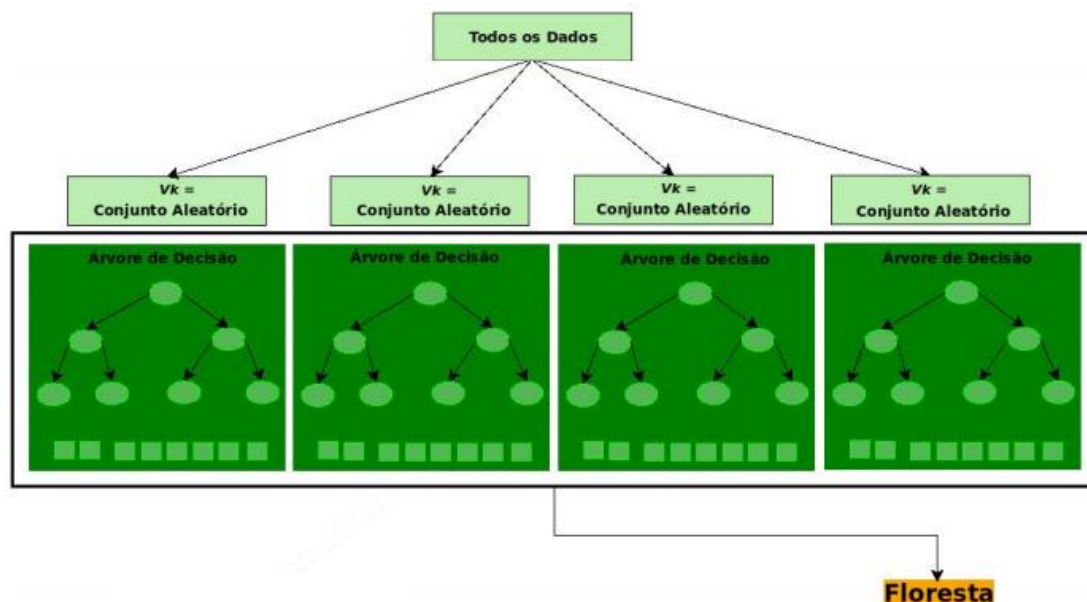
- Flexibilidade: por ser um método não paramétrico, nenhuma distribuição para os dados é assumida. O espaço de classes é dividido em subespaços, e cada um é ajustado por diferentes modelos. Uma árvore fornece uma cobertura exaustiva do subespaço de classes;
- Robustez: possui baixa sensibilidade a distribuições com grande número de valores atípicos;

- Seleção de atributos: o processo de construção de uma árvore seleciona os atributos mais importantes para o modelo de decisão, criando modelos que tendem a ser bastante robustos;
- Interpretabilidade: Decisões complexas e globais podem ser aproximadas por decisões locais mais simples e de fácil visualização.

2.5.2 *Random forest*

Uma *random forest* pode ser descrita como um classificador formado por um conjunto de árvores de decisão, em que conjuntos aleatórios de variáveis são amostrados de forma independentes e distribuídos igualmente em todas as árvores da floresta, formando a entrada de dados de cada uma das árvores. O resultado do processo de classificação é obtido como a classe com o maior número de votos entre as árvores consideradas. Essa técnica permite que se obtenha modelos muito eficazes sem nenhuma preparação dos dados ou conhecimento de modelagem (BREIMAN; CUTLER 2014). A Figura 6 exemplifica uma *random forest*. Nela os atributos de entrada de cada árvore de decisão são definidos pelos vetores aleatórios V_k .

Figura 6 – Exemplo de *random forest*



Fonte: Ibañez, 2016

Os passos para o desenvolvimento da *random forest* são: realizar a seleção do conjunto de dados para o crescimento das árvores, selecionar as características dos

preditores, o crescimento da árvore, o qual é executado até o atendimento de um critério de parada e, por fim a estimativa do erro de crescimento da árvore (IBAÑEZ, 2016).

O algoritmo possui a capacidade de identificar complexas interações entre os dados. Isso permite que, em conjunto, as árvores de decisão possuam uma capacidade de generalização que explique o comportamento dos dados, ao mesmo tempo que evita uma memorização do modelo (BREIMAN et al. 1984). Isso resulta em melhorias significativas na precisão da classificação, que é realizada a partir de um sistema de votação da classe mais popular entre as árvores criadas (BREIMAN, 2001).

2.6 Agrupamento de dados

O objetivo de uma técnica de agrupamento é encontrar uma estrutura de grupos (*clusters*) nos dados, em que os objetos pertencentes a cada grupo compartilhem alguma característica ou propriedade, ou seja, são de alguma forma similares (JAIN; DUBES; CHEN, 1987). É uma técnica não supervisionada de dados pois a similaridade entre os atributos é uma característica intrínseca por não precisar de um arquivo de treinamento com classes pré-definidas (KRONBAUER; FONTOURA; WINCK, 2016).

Faceli et al. (2001) divide os algoritmos de agrupamento de acordo com o método adotado para definir os grupos, classificando os algoritmos em quatro tipos:

- Hierárquicos: geram, a partir de uma matriz de proximidade, uma sequência de partições aninhadas. Suas soluções são tipicamente representadas por um dendograma, consistindo numa árvore binária que representa uma hierarquia de partições;
- Particionais: otimizam o critério de agrupamento utilizando uma técnica iterativa. Isso é feito através da criação de uma partição inicial onde em seguida os objetos são movidos entre os *clusters* com o objetivo de melhorar o agrupamento;
- Baseados em densidade: nesses se assume que os *clusters* são regiões de alta densidade de objetos, separadas por regiões de baixa densidade. Esses algoritmos são capazes de obter *clusters* de formas arbitrárias;

- Baseados em grade: esse grupo de algoritmos define uma grade para o espaço de dados e realiza as operações sobre esse espaço.

Agrupamento de dados pode ser formulado como um problema de otimização com múltiplos objetivos, em que o algoritmo a se escolher e seus parâmetros (valores como a função de distância, o limiar de densidade ou o número esperado de grupos) dependem dos dados e do tipo de resultado procurado (KRONBAUER; FONTOURA; WINCK, 2016).

2.6.1 K-médias (*k-means*)

É o principal representante dos algoritmos de agrupamento particionais. Esse algoritmo particiona o conjunto de dados em k grupos, em que o valor de k é fornecido pelo usuário (DUDA; HART; STORK, 2001).

O algoritmo realiza uma busca para determinar um ponto que represente cada uma das partições, esse ponto é o centro de massa da partição (centroide). Após essa busca, os pontos são comparados com o centroide e agrupados com a partição de maior similaridade. Depois de todos os pontos estarem agrupados, o centroide é recalculado e o processo se repete até atingir determinado critério de parada (KRONBAUER; FONTOURA; WINCK, 2016).

O objetivo desse algoritmo é obter uma partição que minimize o erro médio quadrático para um número k de grupos. O erro quadrático é dado pela Equação (38), em que $\bar{x}^{(j)}$ é o centroide do grupo G_j e $d(x_i, \bar{x}^{(j)})$ é uma medida de distância entre um objeto x_i e o centroide $\bar{x}^{(j)}$.

$$E = \sum_{j=1}^k \sum_{x_i \in G_j} d(x_i, \bar{x}^{(j)})^2 \quad (38)$$

Dentre as principais métricas de distância se destacam a distância euclidiana, que representa a distância entre dois pontos num espaço n -dimensional, a distância de Mahalanobis, que baseada nas correlações entre observações, e a distância de Hamming, que analisa a distância entre dois vetores binários calculando o menor número de substituições necessárias para transformar um vetor no outro.

2.6.2 Validação do número de grupos

Técnicas para validação de grupos são utilizadas para avaliar os resultados de um determinado algoritmo de agrupamento (SOUSA; BARRETO; MEDEIROS, 2017).

Determinar o número ótimo de grupos em um conjunto de dados é uma questão fundamental quando se utiliza algoritmos particionais de agrupamento. A escolha do número ótimo de grupos é subjuntiva e depende do método utilizado para medir semelhanças entre os dados e os parâmetros utilizados no algoritmo (KASSAMBARA, 2015).

Kassambara (2015) divide em dois os grupos de metodologias mais usuais para escolha do número de grupos, são eles:

- Métodos diretos: realizam a otimização de um critério, como a soma do erro quadrático entre grupos ou a silhueta média;
- Métodos de teste: consistem em um teste de hipóteses realizado sobre o resultado do agrupamento.

Um dos métodos diretos e simples de se aplicar é o índice de Davies-Bouldin. Proposto por Davies e Bouldin (1979), esse índice é uma função da relação entre a soma da dispersão dentro do grupo e a separação entre grupos. Para o cálculo do índice, deve ser calculada a dispersão dentro do i -ésimo grupo e a separação do mesmo com o j -ésimo grupo, isso é realizado conforme a equação (39)

$$S_i = \left[\frac{1}{n_i} \sum_{l=1}^{n_i} d(x_l^{(i)}, \bar{x}^{(i)})^2 \right]^{\frac{1}{2}} \quad (39)$$

sendo $x_l^{(i)}$ a observação l no grupo i e S_i a dispersão no grupo i . O índice de Davies-Bouldin pode ser calculado para um número k de grupos como:

$$DB(k) = \frac{1}{k} \sum_{i=1}^k R_i \quad (40)$$

onde R_i é a separação entre grupos calculada por

$$R_i = \max_{j \neq i} \left\{ \frac{S_i + S_j}{d_{ij}} \right\} \quad (41)$$

e d_{ij} é a distância entre os centroides dos grupos i e j .

O cálculo do índice é realizado até um número máximo de grupos definidos pelo usuário e o número de grupos é escolhido como aquele que possui o menor valor do índice calculado.

2.7 Visualização de dados em alta dimensão

Devido ao aumento da complexidade de processos industriais e da elevada automação dos mesmos, os dados coletados em aplicações industriais estão crescendo de forma constante, tanto na quantidade de dados coletados quanto no número de variáveis medidas ou simuladas em um único instante, o que acarreta numa maior complexidade ao lidar com esses conjuntos de dados. Um método comumente escolhido em aplicações como classificação ou visualização é reduzir o número de dimensões dos dados (ENGEL; HÜTTENBERGER; HAMANN, 2012).

As técnicas de redução de dimensão têm como objetivo calcular um conjunto de dados com menor dimensão, que ainda represente padrões e características dos dados originais. Para fins de visualização, as abordagens geralmente projetam o espaço de alta dimensão para um espaço de duas ou três dimensões que pode ser visualizado usando gráficos de dispersão. Um grande número de técnicas de redução da dimensionalidade que visam preservar a estrutura local dos dados foram propostas. Entre as mais utilizadas estão uso das duas ou três componentes principais da matriz de carregamento obtida com o uso da PCA e o uso de Sammon *plots* (SAMMON, 1969).

Apesar do bom desempenho dessas técnicas em conjuntos de dados artificiais, muitas vezes eles não são bem sucedidos na visualização de alta dimensão. Em particular, a maioria das técnicas não é capaz de manter a estrutura local e global dos dados em um único espaço (MAATEN; HINTON, 2008).

Para melhorar a representação gráfica de dados em alta dimensão é utilizado o *t - Distributed Stochastic Neighbor Embedding* (t-SNE). Uma técnica de redução de dimensionalidade para a visualização de dados com alta dimensão proposta por van der Maaten e Hinton (2008). Essa técnica não-linear de redução de dimensionalidade é adequada para transformar dados com muitos atributos para uma espaço bi ou tridimensional, onde se torna possível visualizar o conjunto de dados em um gráfico de

dispersão. Esse algoritmo cria um modelo em um gráfico de dispersão onde cada elemento do conjunto é um ponto, a proximidade de dois pontos indicam que estes são similares, assim como pontos distantes são dissemelhantes (MAATEN; HINTON, 2008).

3 Metodologia proposta

A metodologia proposta utiliza dados de falhas já conhecidas para treinar e aplicar classificadores locais para diagnóstico de novas falhas. Cada classificador é treinado para um subconjunto de falhas que possuam similaridade nas variáveis afetadas após a ocorrência de uma falha.

Inicialmente, ocorrências de falhas conhecidas são agrupadas de forma não supervisionada utilizando o vetor de direção de falha, referente a cada falha. Nesse passo é possível que um mesmo tipo de falha esteja alocado em mais de um grupo. Devido a ocorrência de ruídos, variáveis que não auxiliaram para falha podem ser detectadas como participantes da falha, o que altera o vetor de direção de falha entre uma amostra e outra de um mesmo tipo de falha. Em seguida, para cada um dos grupos calculados é treinado um classificador especializado, que é capaz de identificar o subconjunto de falhas alocadas em um grupo.

Eventualmente, é possível haver grupos com uma única falha, o que dispensa o uso de classificadores. Quando uma falha for detectada, inicialmente ela será associada em um dos grupos existentes e, se no grupo houver somente um tipo de falha o diagnóstico é automático, caso contrário o classificador responsável pelo grupo é aplicado sobre os dados da nova falha, e irá classificá-la entre uma daquelas para que o classificador foi treinado.

Para realizar o diagnóstico, é necessário considerar o comportamento do processo após uma falha. Para isso, a metodologia proposta aplica uma extração de características sobre os dados escolhidos para realização do diagnóstico. A aplicação da metodologia não é limitada a um único tipo de dado de entrada. Nesse trabalho, diferentes tipos de dados são utilizados afim de fornecer um comparativo a respeito de qual deles fornece os melhores resultados de diagnósticos.

A aplicação da proposta é realizada no Capítulo 4, onde foram utilizados estudos de caso para comparar a aplicação de um classificador único com o uso de múltiplos classificadores locais.

Os estudos de casos apresentados no próximo capítulo são realizados utilizando as medidas das variáveis dos processos analisados, os valores de contribuição (seção 2.2.1) depois de detectada uma falha e, em casos onde o processo estudado pode operar em mais

de uma região de operação é proposto o uso dos resíduos calculados a partir da técnica de análise externa (seção 2.3) , o uso dos desvios das variáveis e o cálculo de contribuições dos modelos baseados em análise externa.

As características são utilizadas para treinar os classificadores locais para o diagnóstico. Nas subseções seguintes são apresentadas em mais detalhes as etapas da metodologia.

3.1 Obtenção dos modelos de PCA e DPCA

Um importante passo na correta utilização da análise por componentes principais está na definição dos dados de treinamento para obtenção do modelo de PCA e DPCA. A quantidade de dados deve ser grande o suficiente para reproduzir a variabilidade natural dos sinais, mas com o cuidado de não incluir trechos sob a influência de uma falha.

Para o cálculo dos limiares de controle é necessário definir o número de componentes principais a serem utilizadas e o nível de significância. Nesse trabalho, foi utilizado para a seleção do número de componentes principais o critério da variância percentual acumulada (seção 2.1.2.1), o número de componentes foi escolhido de forma a representar pelo menos 90% da variância total do sistema, esse valor é o mais comum utilizado na literatura (LIU; CHEN, 2014). Para o cálculo dos limiares estatísticos o nível de significância para a construção do modelo foi de 1%.

No caso da aplicação da DPCA os parâmetros selecionados são os mesmos e o número de atrasos é calculado utilizando o método do valor singular chave, segundo apresentado na seção 2.1.2.3.

3.2 Detecção de falha

Uma vez construído o modelo estatístico multivariado a ser utilizado, é possível acompanhar o comportamento dos sinais através do índice estatístico combinado φ (seção 2.1.2), de forma a monitorar o processo utilizando o limiar de controle calculado.

Para cada nova amostra de dados é realizada a sua normalização, utilizando as médias e desvios padrões do modelo, e calculado o índice combinado. Pontos que ultrapassem o limiar de controle da estatística são considerados sob o efeito de alguma falha no processo.

3.3 Agrupamento de falhas

Uma vez detectada uma falha é possível encontrar quais das variáveis influenciaram para o mau funcionamento do processo a partir do vetor de direção de falha ξ (seção 2.1.2.4). Ao analisar a ocorrência de diversas falhas (p falhas) se torna possível criar uma matriz de direção de falha indicando para cada falha que variáveis foram influenciadas. Caso se tenha mais de um evento de uma mesma falha (N eventos) é possível calcular a matriz de direção de falha para todas as falhas da forma $\Xi \in \mathbb{R}^{N \times p \times m}$.

A aplicação de algoritmos de agrupamento permite então separar falhas que influenciaram nas mesmas variáveis, simplificando o trabalho de diagnóstico. Uma vez que todos os eventos de anormalidade estejam representados na forma de vetores de direção de falha, um algoritmo de aprendizado não supervisionado é aplicado para obter os grupos no espaço binário multidimensional definido por vetores $\xi \in \{0,1\}^m$, indicando que o vetor de direção de falhas é formado por zeros e uns e tem tamanho m .

Nesse trabalho, é utilizado o algoritmo de agrupamento k-médias (seção 2.6.1) para calcular os grupos aos quais as falhas serão atribuídas. A seleção do número de grupos é realizada a partir do critério de Davies-Bouldin (seção 2.6.2) onde o número máximo de grupos será o número de falhas conhecidas utilizadas no treinamento dos classificadores. A métrica mais usual para esse algoritmo é a distância Euclidiana, porém ela não é aplicável para vetores binários como o vetor de direção de falhas, sendo assim é utilizada a distância de Hamming, que representa uma medida de similaridade para vetores binários (NOROUZI et al., 2012).

3.4 Dados para os classificadores

Utilizando o conhecimento contido em dados de falhas rotuladas previamente, por especialistas no processo estudado, é possível treinar os classificadores para que seja realizado o diagnóstico no caso de um novo evento não rotulado de falha. Uma vez selecionado o tipo de dado que se deseja analisar, é obtida uma matriz D^j contendo um intervalo de dados do instante de detecção da falha (t) a f instantes após a detecção para j -ésima falha.

Antes de treinar e usar um classificador local para um determinado grupo g , é extraído um conjunto de onze características da matriz de D^j (seção 2.3) para auxiliar no trabalho

de classificação. Nessa etapa, apenas os dados históricos de falhas que estão presentes no grupo analisado são utilizados. Para isso, os dados dessa matriz são divididos em intervalos iguais de d observações, sobre o qual será extraído um conjunto de características de cada uma das m variáveis. Para realizar a extração de características, d deve ser escolhido de forma que ele seja múltiplo do intervalo de dados após a falha, f .

O conjunto de características extraídas em cada intervalo forma um vetor de características que será utilizado pelo classificador. Unindo-se esses vetores é possível obter a matriz de características referentes a j -ésima falha presente no grupo g como $C_H^j \in \mathbb{R}^{\left(\frac{f}{d} \times 11m\right)}$. Nessa etapa, o intervalo para extração de características d deve ser múltiplo do intervalo de dados em falha f para compor a matriz C_H^j .

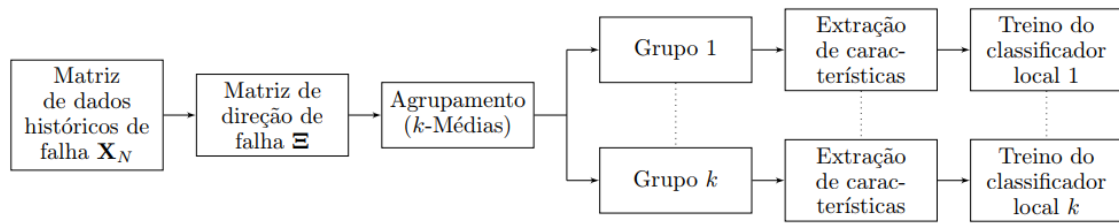
3.5 Treinamento dos classificadores

A matriz de características é utilizada com o objetivo de treinar os classificadores específicos. Cada classificador específico a um grupo é treinado apenas com dados de falhas associadas a tal grupo. Nesse trabalho, são utilizados classificadores do tipo *random forest* (seção 2.5.2). Esse classificador é robusto e possui elevado desempenho mesmo com sinais ruidosos. O objetivo desse classificador é criar um modelo capaz de prever o rótulo da falha utilizando regras de decisão inferidas das matrizes de características.

A Figura 7 ilustra a metodologia da etapa de treinamento dos classificadores para diagnóstico de falha. Nela X_N representa uma matriz de dados históricos de N falhas em que o diagnóstico já foi realizado e, portanto, as amostras já estão rotuladas com as falhas correspondentes. Essa matriz que é utilizada para que sejam obtidos os vetores de direção de falha de cada um dos eventos contidos na mesma, a união dos vetores de direção de falha gera a matriz de direção de falha Ξ .

O algoritmo de agrupamento k-médias é aplicado e são calculados os centroides e as falhas contidas nos k grupos. Para cada grupo é obtida uma matriz de características, utilizando apenas as falhas associadas ao grupo, que é utilizada no treinamento do classificador.

Figura 7 – Treinamento dos classificadores



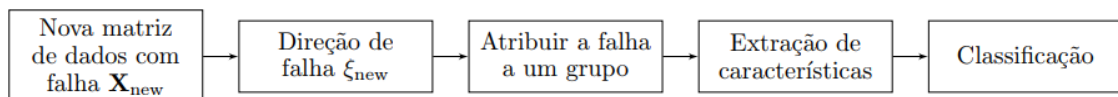
Fonte: O próprio autor

Uma possibilidade seria incluir, durante a etapa de treinamento, dados em funcionamento normal do sistema. Isso poderia mitigar o efeito de falsos alarmes, pois essas ocorrências seriam classificadas como situação de normalidade. Um problema dessa abordagem ocorre quando a falha possui uma pequena intensidade, sendo classificada de forma incorreta como funcionamento normal. Nesse caso um evento de falha será erroneamente classificado como um falso positivo.

3.6 Aplicação

Após o agrupamento das falhas e treinamento dos classificadores locais, o modelo pode ser usado para o diagnóstico de falhas. A Figura 8 ilustra o processo para análise de um novo evento de detecção de falha.

Figura 8 – Aplicação dos classificadores.



Fonte: O próprio autor

Para uma nova matriz de dados em falha X_{new} é extraído o vetor de direção de falha ξ_{new} e, a partir da medida da distância de Hamming para os centroides dos grupos, a nova falha é associada ao grupo mais próximo. Em seguida as características são extraídas e o classificador realiza o diagnóstico da falha.

4 Estudos de caso

Nesse capítulo a metodologia proposta será testada em três diferentes estudos de caso. No primeiro estudo de caso a metodologia é avaliada em sua capacidade de classificar corretamente séries temporais, usando para isto 39 séries temporais disponíveis na literatura para esta finalidade.

Uma vez demonstrada a capacidade de classificação da metodologia em bases de dados conhecidas, a mesma é aplicada para diagnosticar falhas em dois simuladores de processos industriais que são *benchmarks* da literatura. O primeiro simulador testado é o do processo Tennessee Eastman (*Tennessee Eastman process* – TEP), no qual será validada a aplicação da metodologia em um ambiente industrial que opera sobre uma única região de operação. Em seguida é utilizado o simulador de um reator tanque agitado continuamente (*continuous stirred-tank reactor* – CSTR) onde a metodologia é aplicada com diferentes métodos de entrada de dados para uma planta com várias regiões de operação.

4.1 Séries temporais da literatura

Li et al. (2016) demonstra empiricamente, através da aplicação de várias transformadas *wavelets* discretas, que a classificação de séries temporais produz uma precisão semelhante para dados compactados e sem tratamento. Também mostraram que, em alguns conjuntos de dados, as *wavelets* podem realmente ajudar na redução de variações ruidosas que deterioram o desempenho das tarefas de classificação.

Os dados¹ incluem dois conjuntos separados *a priori*, permitindo que resultados disponíveis na literatura sejam comparados com o obtido a partir da aplicação da metodologia proposta.

A Tabela 2 lista as bases de dados utilizadas, o número de classes presente em cada base, o número de séries temporais utilizadas para treinamento e validação do classificador, e o número de pontos de cada série temporal.

¹ Todas as bases de dados estão disponíveis em <https://github.com/serval-snt-uni-lu/wavelets-tsc>.

Tabela 2 – Bases de dados

Base de dados	Classes	Treinamento	Teste	Nº de pontos
ArrowHead	3	36	175	251
BeetleFly	2	20	20	512
BirdChicken	2	20	20	512
Computers	2	250	250	720
DistalPhalanxOutlineAgeGroup	3	139	400	80
DistalPhalanxOutlineCorrect	2	276	600	80
DistalPhalanxTW	6	139	400	80
Earthquakes	2	139	322	512
ECG5000	5	500	4500	140
ElectricDevices	7	8926	7711	96
FordA	2	1320	3601	500
FordB	2	810	3636	500
Ham	2	109	105	431
HandOutlines	2	370	1000	2709
Herring	2	64	64	512
InsectWingbeatSound	11	220	1980	256
LargeKitchenAppliances	3	375	375	720
Meat	3	60	60	448
MiddlePhalanxOutlineAgeGroup	3	154	400	80
MiddlePhalanxOutlineCorrect	2	291	600	80
MiddlePhalanxTW	6	154	399	80
PhalangesOutlinesCorrect	2	1800	858	80
Phoneme	39	214	1896	1024
ProximalPhalanxOutlineAgeGroup	3	400	205	80
ProximalPhalanxOutlineCorrect	2	600	291	80
ProximalPhalanxTW	6	205	400	80
RefrigerationDevices	3	375	375	720
ScreenType	3	375	375	720
ShapeletSim	2	20	180	500
ShapesAll	60	600	600	512

SmallKitchenAppliances	3	375	375	720
Strawberry	2	370	613	235
ToeSegmentation1	2	40	228	277
ToeSegmentation2	2	36	130	343
UWaveGestureLibraryAll	8	896	3582	945
Wine	2	57	54	234
WordSynonyms	25	267	638	270
Worms	5	77	181	900
WormsTwoClass	2	77	181	900

Fonte: Adaptado de Li et al. (2016)

Como as bases de dados são constituídas por séries temporais univariadas ($m = 1$), algoritmos de agrupamento para uso de classificadores locais não são aplicáveis, uma vez que não existem vetores de direção de falha. A extração de características é realizada de forma que cada série temporal seja dividida em cinco intervalos iguais e o classificador *random forest* é treinado. Foram testados diferentes valores para o número de árvores e se percebeu que, a partir da aplicação com 10 árvores de decisão, não houve ganho da métrica de acurácia.

Para avaliar a metodologia são realizadas duas comparações. Inicialmente as acurácias de classificação obtidas utilizando a metodologia proposta são comparadas com a melhor e pior acurácia obtida em Li et al. (2016), esse resultado é ilustrado na Figura 9. A Figura 10 compara os resultados de acurácia da metodologia, com a classificação que obteve a melhor acurácia média em Li et al. (2016). Todas as acurácias citadas anteriormente, bem como as acurácias obtidas através da aplicação da metodologia proposta são apresentadas na Tabela 3.

Tabela 3 – Comparação dos resultados

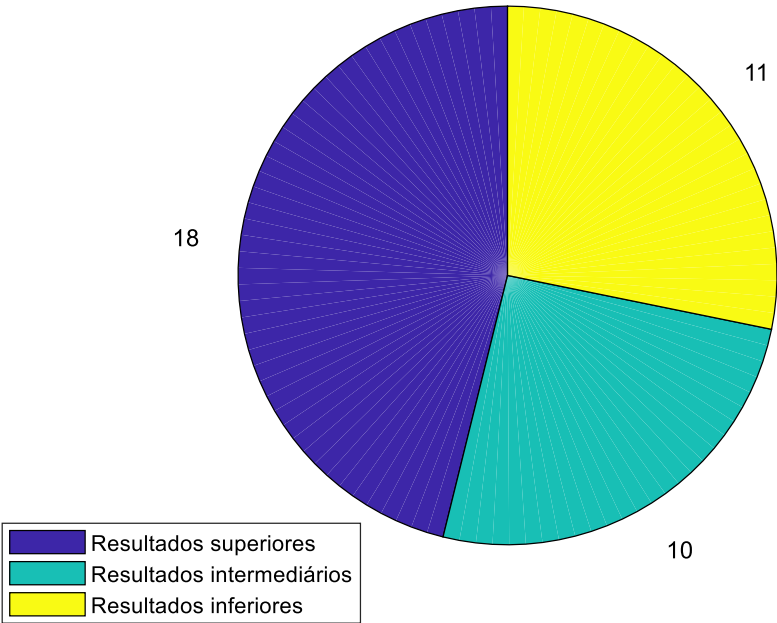
Base de dados	Metodologia proposta	Melhor resultado Li et al. (2016)	Pior resultado Li et al. (2016)	Maior média Li et al. (2016)
ArrowHead	72,9	78,9	69,7	77,7
BeetleFly	69,1	80	65	70
BirdChicken	87,1	75	55	60

Computers	81,5	58,8	47,2	50,4
DistalPhalanxOutlineAgeGroup	70,7	78,5	74,5	75,5
DistalPhalanxOutlineCorrect	73,9	75,3	73	75,2
DistalPhalanxTW	71,7	73,2	67,2	69
Earthquakes	87	75,5	69,6	69,6
ECG5000	94,1	93,3	91,3	92,9
ElectricDevices	58,2	65,6	57,1	60,1
FordA	76	58,2	57	58,2
FordB	65	55,1	53,8	55
Ham	57,9	65,7	50,5	65,7
HandOutlines	74,5	78,8	74,7	78,4
Herring	58,5	56,2	45,3	50
InsectWingbeatSound	57,5	48,3	41	43,6
LargeKitchenAppliances	61	61,1	58,9	59,5
Meat	82,3	93,3	86,7	90
MiddlePhalanxOutlineAgeGroup	57,7	75,8	72,8	75
MiddlePhalanxOutlineCorrect	70,1	74,3	64,5	69,5
MiddlePhalanxTW	57,7	59,4	57,1	58,4
PhalangesOutlinesCorrect	67,9	75,1	71,3	71,4
Phoneme	57,3	10,5	8,6	9,1
ProximalPhalanxOutlineAgeGroup	77	80	75,1	75,6
ProximalPhalanxOutlineCorrect	74,9	79,7	75,3	75,3
ProximalPhalanxTW	75,6	73,8	69,8	72,8
RefrigerationDevices	57,1	48,8	45,9	46,1
ScreenType	57,1	41,1	36,5	37,6
ShapeletSim	57,1	54,4	50,6	51,1
ShapesAll	57,4	68,5	56,2	67,3
SmallKitchenAppliances	71,1	59,2	50,7	55,7
Strawberry	86,8	92,7	91,2	91,8
ToeSegmentation1	71,1	67,5	58,3	65,8
ToeSegmentation2	60,7	83,8	71,5	83,8
UWaveGestureLibraryAll	59,2	93,1	88,1	91,6
Wine	58,9	70,4	61,1	66,7

WordSynonyms	57,9	54,5	37,5	53,9
Worms	68,8	37	28,7	35,4
WormsTwoClass	77,2	61,3	54,7	59,1

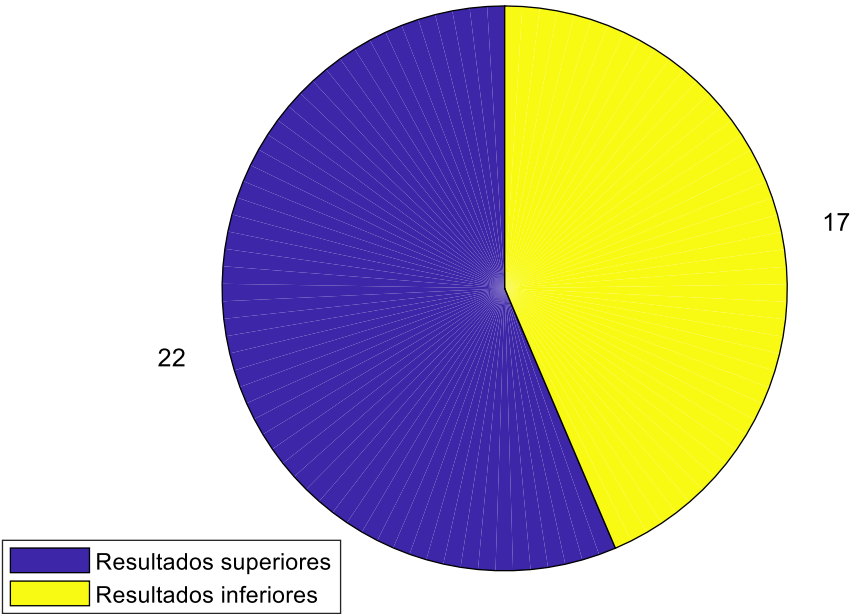
Fonte: O próprio autor

Figura 9 – Comparação com os resultados de Li et. al. (2016).



Fonte: O próprio autor

Figura 10 – Comparação a metodologia de maior média de Li et. al. (2016).



Fonte: O próprio autor

Das 39 séries temporais analisadas, em 18 a metodologia proposta obteve uma acurácia superior quando comparadas ao melhor resultado de Li et al. (2016). Em 10 casos a acurácia esteve no intervalo entre o pior e melhor resultados e houve 11 séries em que a metodologia proposta obteve uma acurácia inferior ao pior de Li et al. (2016). Quando comparadas apenas a metodologia de maior média de Li et al. (2016), as acurácias da metodologia proposta apresentaram melhores resultados em 22 das 39 séries temporais.

Os resultados apresentados nessa seção demonstram que a metodologia proposta, utilizando extração de características e um classificador *random forest*, gera bons resultados, quando comparada a literatura apresentada, para classificação de séries temporais univariadas. Com isso, propõe-se aplicar a metodologia para diagnosticar falhas em processos industriais, através da classificação de séries temporais multivariadas.

4.2 Processo Tennessee Eastman (TEP)

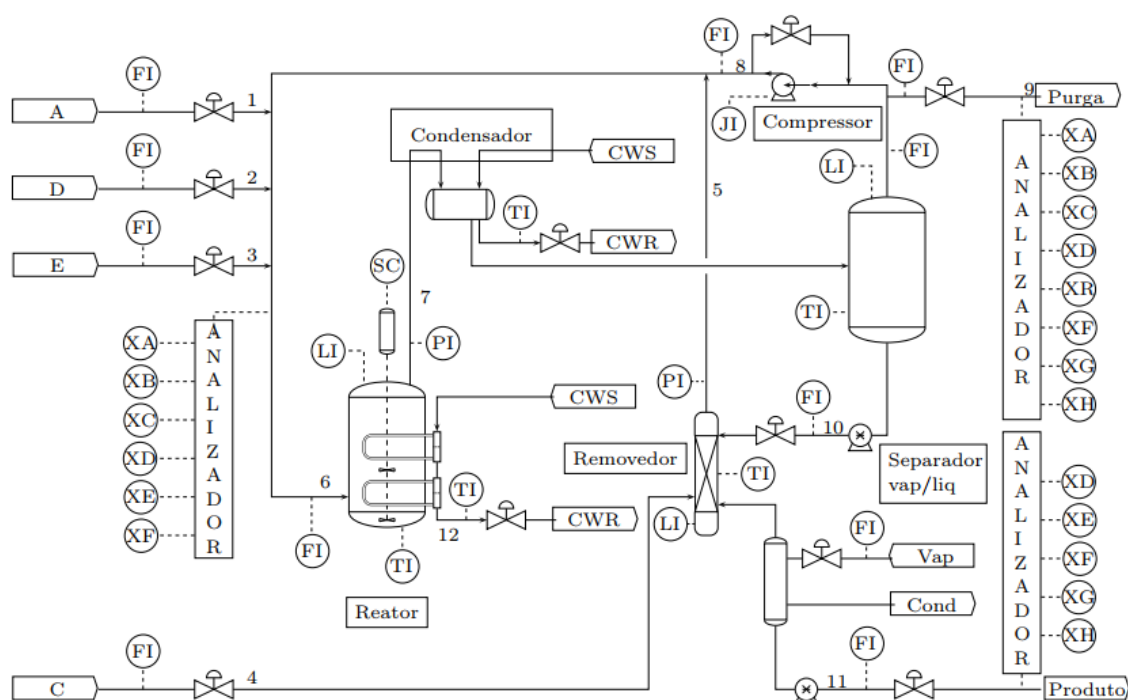
Em um esforço para fornecer um teste realista para tarefas de monitoramento de processos industriais, um simulador de uma planta de produção foi proposto em Downs e Vogel (1993). Devido à complexidade do simulador e a sua proximidade com um processo real, o TEP se tornou um dos *benchmarks* mais difundidos e utilizados na literatura no uso de metodologias de detecção e diagnóstico de falhas. A metodologia é aplicada a esse simulador para analisar o correto diagnóstico das falhas disponíveis e comparar com resultados existentes na literatura atual.

4.2.1 Descrição do simulador

Um reator, condensador, removedor, compressor e separador constituem os principais componentes do sistema. Nele são obtidos dois produtos líquidos (G, H) e um subproduto líquido (F) em duas reações paralelas entre os reagentes A, B, C e D . Todas as reações são irreversíveis e exotérmicas.

Os reagentes alimentam o reator onde ocorre a formação dos produtos líquidos. O reator possuiu uma jaqueta de refrigeração para remover o calor da reação. Componentes não condensados são reciclados através de um compressor centrífugo e retornam à alimentação do reator. Componentes condensados vão para o removedor onde os produtos que são separados. O subproduto é eliminado do sistema no separador.

Figura 11 – Processo Tennessee Eastman



Fonte: Adaptado de Downs e Vogel (1993)

Nas simulações, o intervalo entre as aquisições de duas amostras consecutivas de dados é 3 minutos. A utilidade do simulador é corroborada por publicações recentes no campo de diagnóstico de falhas, como D'Angelo et al. (2016) e Gao e Hou (2016). A Figura 11 ilustra um esquemático da planta simulada. O simulador utiliza 12 variáveis manipuladas (Tabela 4) e 41 variáveis medidas de processo (Tabelas 4 a 7).

Tabela 4 – Variáveis manipuladas do processo

Identificação	Descrição da variável
XMV (1)	Vazão de reagente <i>D</i> (fluxo 2)
XMV (2)	Vazão de subproduto (fluxo 3)
XMV (3)	Vazão de reagente <i>A</i> (fluxo 1)
XMV (4)	Vazão dos reagentes <i>A</i> e <i>C</i> (fluxo 4)
XMV (5)	Válvula do compressor
XMV (6)	Válvula de purga (fluxo 9)
XMV (7)	Vazão de líquido para o separador (fluxo 10)
XMV (8)	Vazão de produto do separador (fluxo 11)
XMV (9)	Válvula de vapor do removedor
XMV (10)	Vazão de refrigerante do reator
XMV (11)	Vazão de refrigerante do condensador
XMV (12)	Velocidade do agitador

Fonte: Adaptado de Downs e Vogel (1993)

Tabela 5 – Variáveis medidas do processo

Identificação	Descrição da variável
XMEAS (1)	Vazão de reagente <i>A</i> (fluxo 1)
XMEAS (2)	Vazão de reagente <i>D</i> (fluxo 2)
XMEAS (3)	Vazão de subproduto (fluxo 3)
XMEAS (4)	Vazão dos reagentes <i>A</i> e <i>C</i> (fluxo 4)
XMEAS (5)	Vazão de reciclagem (fluxo 8)
XMEAS (6)	Alimentação do reator (fluxo 6)
XMEAS (7)	Pressão do reator
XMEAS (8)	Nível do reator
XMEAS (9)	Temperatura do reator
XMEAS (10)	Taxa de purga (fluxo 9)
XMEAS (11)	Temperatura do separador
XMEAS (12)	Nível do separador
XMEAS (13)	Pressão do separador
XMEAS (14)	Fluxo do separador (fluxo 10)
XMEAS (15)	Nível do removedor
XMEAS (16)	Pressão do removedor
XMEAS (17)	Fluxo do removedor (fluxo 11)
XMEAS (18)	Temperatura do removedor
XMEAS (19)	Vazão de fás do removedor
XMEAS (20)	Pressão do compressor
XMEAS (21)	Temperatura de saída do refrigerante do reator
XMEAS (22)	Temperatura de saída do refrigerante do separador

Fonte: Adaptado de Downs e Vogel (1993)

Tabela 6 – Variáveis medidas no fluxo de alimentação

Identificação	Análise do fluxo de alimentação (fluxo 6)
XMEAS (23)	Concentração de <i>A</i>
XMEAS (24)	Concentração de <i>B</i>
XMEAS (25)	Concentração de <i>C</i>
XMEAS (26)	Concentração de <i>D</i>
XMEAS (27)	Concentração de <i>E</i>
XMEAS (28)	Concentração de <i>F</i>

Fonte: Adaptado de Downs e Vogel (1993)

Tabela 7 – Variáveis medidas no gás de purga

Identificação	Análise do gás de purga (fluxo 9)
XMEAS (29)	Concentração de <i>A</i>
XMEAS (30)	Concentração de <i>B</i>
XMEAS (31)	Concentração de <i>C</i>
XMEAS (32)	Concentração de <i>D</i>
XMEAS (33)	Concentração de <i>E</i>
XMEAS (34)	Concentração de <i>F</i>
XMEAS (35)	Concentração de <i>G</i>
XMEAS (36)	Concentração de <i>H</i>

Fonte: Adaptado de Downs e Vogel (1993)

Tabela 8 – Variáveis medidas no fluxo de produto

Identificação	Análise do produto (fluxo 11)
XMEAS (37)	Concentração de <i>D</i>
XMEAS (38)	Concentração de <i>E</i>
XMEAS (39)	Concentração de <i>F</i>
XMEAS (40)	Concentração de <i>G</i>
XMEAS (41)	Concentração de <i>H</i>

Fonte: Adaptado de Downs e Vogel (1993)

Todas as variáveis manipuladas e medidas são incorporadas nos vetores de dados, de dimensão 53, que reflete a descrição completa do estado do processo. As falhas disponíveis no simulador estão listadas na Tabela 9. Originalmente, 20 falhas foram definidas em Downs e Vogel (1993) e uma falha adicional de agarramento de válvula foi definida em Chiang et al. (2002).

Tabela 9 – Falhas do processo

Identificação	Descrição da variável	Tipo de variação
IDV (1)	Razão de alimentação <i>A/C</i> (fluxo 4)	Degrau
IDV (2)	Composição de <i>B</i> (fluxo 4)	Degrau
IDV (3)	Temperatura de alimentação de <i>D</i> (fluxo 2)	Degrau
IDV (4)	Temperatura de entrada do refrigerante do reator	Degrau
IDV (5)	Temperatura de entrada do refrigerante do condensador	Degrau
IDV (6)	Perda de alimentação de <i>A</i> (fluxo 1)	Degrau
IDV (7)	Redução na disponibilidade de <i>C</i> (fluxo 4)	Degrau
IDV (8)	Composição de alimentação de <i>A</i> , <i>B</i> e <i>C</i> (fluxo 4)	Variação aleatória
IDV (9)	Temperatura de alimentação de <i>D</i> (fluxo 2)	Variação aleatória
IDV (10)	Temperatura de alimentação de <i>C</i> (fluxo 4)	Variação aleatória
IDV (11)	Temperatura de entrada do refrigerante do reator	Variação aleatória
IDV (12)	Temperatura de entrada do refrigerante do condensador	Variação aleatória
IDV (13)	Cinética das reações	Afastamento lento
IDV (14)	Válvula do refrigerante do reator	Agarramento
IDV (15)	Válvula do refrigerante do condensador	Agarramento
IDV (16)	Desconhecido	Desconhecida
IDV (17)	Desconhecido	Desconhecida
IDV (18)	Desconhecido	Desconhecida
IDV (19)	Desconhecido	Desconhecida
IDV (20)	Desconhecido	Desconhecida
IDV (21)	Agarramento de válvula	Agarramento

Fonte: Adaptado de Chiang, Russel e Bratz (2002)

Para avaliar a metodologia dois testes são realizados. Inicialmente a metodologia proposta é comparada com os resultados de diagnósticos obtidos por Gao e Hou (2016). Os autores utilizaram a técnica de máquinas de vetores de suporte (SVM) para realizar o diagnóstico de um subconjunto de falhas do simulador, e obtiveram uma precisão de 96,77% de diagnósticos corretos. Em seguida, a metodologia é aplicada para todas as falhas disponibilizadas pelo simulador.

4.2.2 Detecção e diagnóstico de falhas

Devido ao comportamento dinâmico do processo, para ambos os casos estudados utilizando o simulador do TEP, foi aplicada a técnica de DPCA. Para o uso do algoritmo KSV (seção 2.1.2.3) foi estipulado um valor máximo de 30 atrasos, e após a execução do algoritmo o número de atrasos selecionado foi 5.

Sempre que uma nova amostra de dados é gerada é realizada a sua normalização, utilizando as médias e desvios padrões do modelo obtido com os dados de treinamento, e é calculado o índice combinado. Pontos que ultrapassem o limiar de controle da estatística são considerados sob o efeito de alguma falha no processo.

A partir da detecção da falha são selecionados os intervalos dados após as falhas. Serão comparados os resultados ao se utilizar diretamente os dados para diagnóstico, com o uso de dados de contribuição.

4.2.3 Teste usando um subconjunto das falhas

Gao e Hou (2016) propõem o uso de SVM para diagnosticar as falhas do simulador do TEP. Um subconjunto de falhas composto pelas falhas IDV (1), IDV (2), IDV (4), IDV (5) é utilizado. Com a intenção de testar a abordagem proposta neste trabalho, foram executadas 100 simulações para cada uma dessas quatro falhas. Durante as primeiras 24 horas (480 amostras), o processo funciona em condições de normalidade, antes que uma condição de falha seja introduzida. Em seguida é calculada a matriz de direção de falha \mathbf{E} composta concatenando os vetores de direção de falha de todas as simulações. O agrupamento de dados utilizando o critério de Davies-Bouldin (seção 2.6.2) indicou o que as falhas devem ser separadas em dois grupos, apresentados na Tabela 10.

Tabela 10 – Resultado do agrupamento para um subconjunto de falhas

Grupo	Falhas
1	1, 2, 5
2	2, 4

Fonte: O próprio autor

Na fase de treinamento, para cada classificador local, 70% dos dados são usados para treinar uma *random forest* com 1000 árvores de decisão e 30% são retidos para validar o classificador. Para o diagnóstico são utilizadas 40 amostras após a detecção de falha de cada evento de falha presente no grupo. A extração de características é realizada a cada 20 amostras, resultando em matrizes de características na forma $C_H^j \in \mathbb{R}^{(2 \times 583)}$.

Inicialmente a metodologia é aplicada aos dados das variáveis após uma falha ser detectada. A Tabela 11 compara as acurácias do classificador de cada grupo e a acurácia de um classificador único quando utilizadas as medições das variáveis como entrada para os classificadores. Analisando as acurácias pode-se concluir que a aplicação de classificadores locais melhora a acurácia global da classificação.

Tabela 11 – Acurácia da metodologia para um subconjunto de falhas utilizando as medições das variáveis

Grupo	Acurácia de cada classificador (%)	Acurácia do classificador único (%)
1	89,63	-
2	85,00	-
Global	87,78	83,15

Fonte: O próprio autor

A Tabela 12 compara as acurácias estimadas utilizando classificadores locais a acurácia de um único classificador para diagnóstico do subconjunto de falhas, quando utilizados como entrada as contribuições das variáveis (seção 2.2.1). Cada classificador local desempenha bem a tarefa de diagnóstico para as falhas pertencentes ao seu grupo. A performance global é 96,00%, consideravelmente maior que a de um único classificador com 91,28%.

Tabela 12 – Acurácia da metodologia para um subconjunto de falhas utilizando as contribuições

Grupo	Acurácia de cada classificador (%)	Acurácia do classificador único (%)
1	100,00	-
2	90,00	-
Global	96,00	91,28

Fonte: O próprio autor

Comparando os resultados obtidos nas Tabelas 10 e 11 é possível perceber a vantagem em se utilizar os dados de contribuição para o diagnóstico das falhas, uma vez que esses apresentam melhores resultados quando comparados ao uso das medições das variáveis.

Após o estágio de treinamento dos classificadores, novas simulações são realizadas para validar seu desempenho. Como a validação que apresentou melhor resultado utilizou as contribuições, esses dados foram utilizados para validar a metodologia com novas simulações. A Figura 12 mostra a matriz de confusão resultante para a classificação e o diagnóstico, quando 100 novas simulações de cada da falha são realizadas. Nas linhas da matriz estão as falhas simuladas e nas colunas as falhas previstas utilizando os classificadores. A diagonal da matriz contém o número de simulações onde a classificação foi realizada corretamente, e a acurácia pode ser calculada como a razão entre a soma da diagonal principal e o número total de simulações.

Figura 12 – Matriz de confusão para o subconjunto de falhas.

Acurácia = 99.25%

Falha simulada	IDV(1)	100	0	0	0
	IDV(2)	0	100	0	0
	IDV(4)	0	3	97	0
	IDV(5)	0	0	0	100
		IDV(1)	IDV(2)	IDV(4)	IDV(5)
		Falha predita			

Fonte: O próprio autor

Conforme mencionado, os resultados da abordagem proposta são comparados ao resultado obtido por Gao e Hou (2016), nesse trabalho os autores utilizam dados do

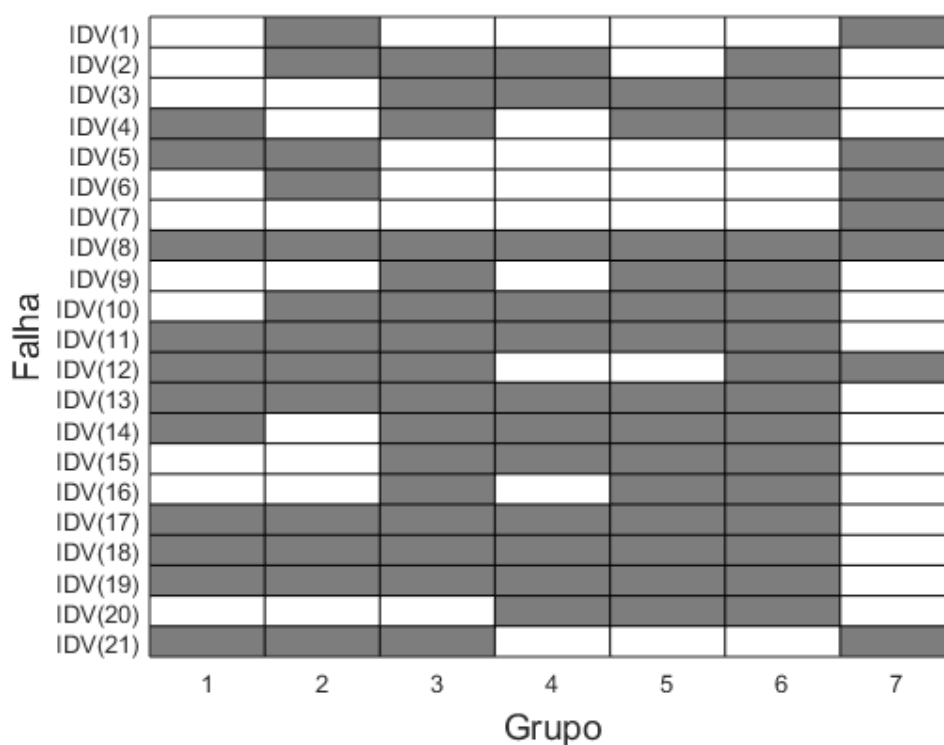
subconjunto apresentado de falhas para treinar o classificador proposto, porém o teste do mesmo é realizado apenas com falhas pertencentes a classe IDV (1). A melhor acurácia obtida pelos autores usando SVM foi de 96,77%.

A abordagem proposta nesta dissertação obteve 100% de acurácia na classificação da falha IDV (1) e ofereceu bons resultados quando aplicadas falhas de todo o conjunto selecionado, obtendo uma acurácia de 99,25%.

4.2.4 Todas as falhas

Os testes anteriores, com um subconjunto de quatro falhas mostraram bons resultados. O próximo passo será testar a metodologia para o conjunto completo de 21 falhas, para corroborar a robustez da mesma em relação a uma gama mais ampla de condições de falha. O conjunto de dados é construído com 100 simulações de cada falha e 24 horas de operação normal antes que as falhas ocorram. O monitoramento do processo é feito novamente usando DPCA com 5 atrasos. A matriz de direção de falha é construída e, utilizando o critério de Davies-Bouldin, sete grupos são definidos. As falhas associadas a cada grupo são apresentadas na Figura 13. Nela os elementos preenchidos indicam a presença da falha no grupo correspondente.

Figura 13 – Falhas presentes em cada grupo



Fonte: O próprio autor

Novamente a metodologia será aplicada utilizando as medições das variáveis e a acurácia global é comparada com a obtida utilizando os dados de contribuições. Um único classificador independente é comparado com os classificadores locais. As acurácias calculadas utilizando as medições das variáveis estão expostas na Tabela 13 e as acurácias utilizando dados de contribuição estão na Tabela 14.

Tabela 13 – Acurácia da metodologia utilizando as medições das variáveis para todas as falhas.

Grupo	Acurácia de cada classificador (%)	Acurácia do classificador único (%)
1	99,17	-
2	94,26	-
3	68,13	-
4	84,44	-
5	63,57	-
6	66,56	-
7	96,44	-
Global	75,85	64,59

Fonte: O próprio autor

Tabela 14 – Acurácia da metodologia utilizando os dados de contribuição para todas as falhas.

Grupo	Acurácia de cada classificador (%)	Acurácia do classificador único (%)
1	98,89	-
2	95,37	-
3	73,33	-
4	88,67	-
5	70,00	-
6	72,44	-
7	97,78	-
Global	83,33	70,35

Fonte: O próprio autor

Comparando os resultados obtidos nas Tabelas 12 e 13, é possível notar que, ao se utilizar os dados de contribuição para o diagnóstico das falhas, são obtidos os melhores resultados.

Para validar a abordagem com novos dados, são realizadas 100 novas simulações para cada falha. Os vetores de direção de falha são calculados e, cada falha é associada a um grupo, onde o classificador local executará o diagnóstico de falha. A matriz de confusão para o conjunto completo de falhas é mostrada na Figura 14.

Figura 14 – Matriz de confusão para o todas as de falhas.

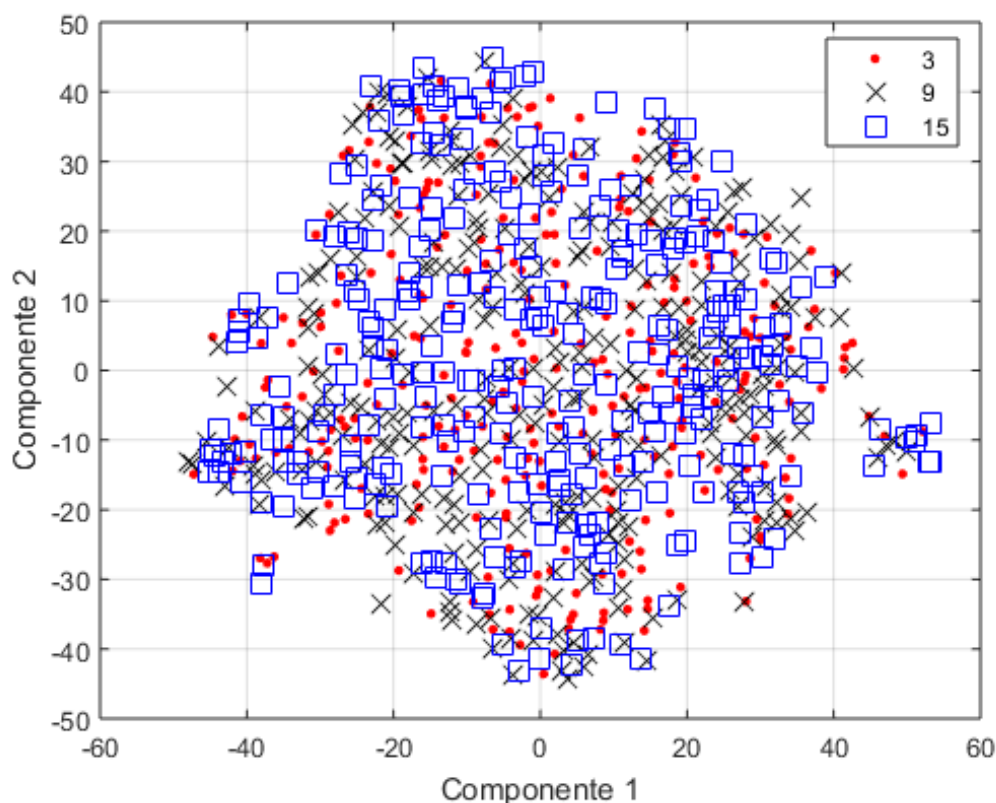
Acurácia = 88.29%

Falha simulada	IDV(1)	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	IDV(2)	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	IDV(3)	0	0	36	0	0	0	0	1	40	0	0	0	0	0	21	1	0	0	1		
	IDV(4)	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	IDV(5)	0	0	0	0	99	0	0	0	0	0	0	0	0	0	1	0	0	0	0		
	IDV(6)	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0		
	IDV(7)	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0		
	IDV(8)	2	0	0	0	0	0	0	97	1	0	0	0	0	0	0	0	0	0	0		
	IDV(9)	0	0	32	0	0	0	0	0	43	6	0	0	1	0	17	1	0	0	0		
	IDV(10)	0	0	3	0	0	0	0	0	2	87	0	0	0	0	2	6	0	0	0		
	IDV(11)	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0		
	IDV(12)	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0		
	IDV(13)	0	0	5	0	0	0	0	0	5	0	0	0	87	0	2	0	0	0	1		
	IDV(14)	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0		
	IDV(15)	0	0	12	0	0	0	0	0	33	3	0	0	1	0	48	3	0	0	0		
	IDV(16)	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	99	0	0	0		
	IDV(17)	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2	0	96	0	0		
	IDV(18)	0	0	9	0	0	0	0	0	6	0	0	3	0	0	11	0	0	71	0		
	IDV(19)	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	99	0		
	IDV(20)	0	0	1	0	0	0	0	0	2	0	0	0	0	0	5	0	0	0	92		
	IDV(21)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
		IDV(1)	IDV(2)	IDV(3)	IDV(4)	IDV(5)	IDV(6)	IDV(7)	IDV(8)	IDV(9)	IDV(10)	IDV(11)	IDV(12)	IDV(13)	IDV(14)	IDV(15)	IDV(16)	IDV(17)	IDV(18)	IDV(19)	IDV(20)	IDV(21)
		Falha predita																				

Fonte: O próprio autor

Em seguida aplica-se a técnica t-SNE para analisar a proximidade entre as características das falhas onde ocorreu a maior taxa de erros, são elas: IDV(3), IDV(9) e IDV(15). A Figura 15 mostra como as características são extremamente similares, o que justifica a alta taxa de erro do classificador. Portanto, a abordagem proposta pode fornecer um diagnóstico errado para essas falhas. Um possível motivo é que o processo reage de forma semelhante a ocorrência dessas falhas, para as características calculadas nesse trabalho.

Figura 15 – Visualização bidimensional das características utilizando t-SNE para as falhas IDV(3), IDV(9) e IDV(15).



Fonte: O próprio autor

Analisando a Tabela 9, é possível perceber que dentre as falhas geradas pelo simulador, existem aquelas que possuem a mesma descrição, porém com o tipo de variação diferente. As falhas equivalentes são IDV(3) e IDV(9), IDV(4) e IDV(11), IDV(5) e IDV(12). De um ponto de vista de identificação de séries temporais ocorrem erros na discriminação dos dados analisados, porém de um ponto de vista de diagnóstico de falha identificar por exemplo a falha IDV(3) como IDV(9) não seria um erro, uma vez que ambas são falhas na mesma variável.

Para ilustrar o sucesso da aplicação da metodologia, o resultado obtido é comparado com três outros resultados apresentados na literatura. Em Jing e Hou (2015) os autores utilizam as técnicas de PCA e SVM para realizar o diagnóstico das falhas. É obtida uma acurácia de 40,53% para o uso de SVM e, 65,09% quando realizado o diagnóstico a partir da maior pontuação obtida com a PCA. Em Eslamloueyan (2010) o autor utiliza DPCA para detecção de falhas, associado com o uso de redes neurais para diagnóstico, obtendo uma acurácia média de 67,92%.

Em D'Angelo et. al. (2016) os autores utilizam lógica nebulosa com algoritmo Metropolis-Hastings para detecção de falha. O diagnóstico é realizado utilizando um algoritmo imune auxiliado por uma rede neural Kohonen para isolar a falha. As bases de dados² utilizadas pelos autores consistem em 42 conjuntos de dados, separados em 21 conjuntos de treinamento e 21 conjuntos teste. Existe uma diferença entre os conjuntos de dados utilizados pelos autores e os conjuntos de dados utilizados nesse trabalho.

D'Angelo et al. (2016) fornecem o resultado da aplicação através da matriz de confusão. Esse resultado foi utilizado para calcular a acurácia do método, que foi de 88,23%. Aplicando o treinamento com várias simulações de cada falha, os resultados obtidos com a metodologia proposta fornecem uma acurácia ligeiramente maior, 88,29%.

A metodologia proposta apresenta resultados melhores que os presentes na literatura citada, utilizando algoritmos de fácil programação e aplicação, quando estudado o simulador *Tennessee Eastman*.

4.3 Reator tanque agitado contínuo

A metodologia proposta é então aplicada no *benchmark* CSTR, proposto por Finch (1989) e melhorado por Oyeleye (1990). Nele é simulado o funcionamento de uma planta química que consiste num reator com resfriamento externo, devido a reações exotérmicas que ocorrem no processo.

Na seção anterior foi aplicada a metodologia para um processo industrial que possui uma única região de operação. No caso do simulador CSTR é possível, através da alteração dos *setpoints* de temperatura e nível do reator, simular a operação da planta em diferentes regiões de operação.

Nesse estudo de caso é aplicada a técnica de análise externa para calcular resíduos em diferentes pontos de operação. Esses resíduos são utilizados para detectar as falhas e calcular os vetores de direção de falha utilizados para o agrupamento.

² As bases de dados estão disponíveis em <http://web.mit.edu/braatzgroup>

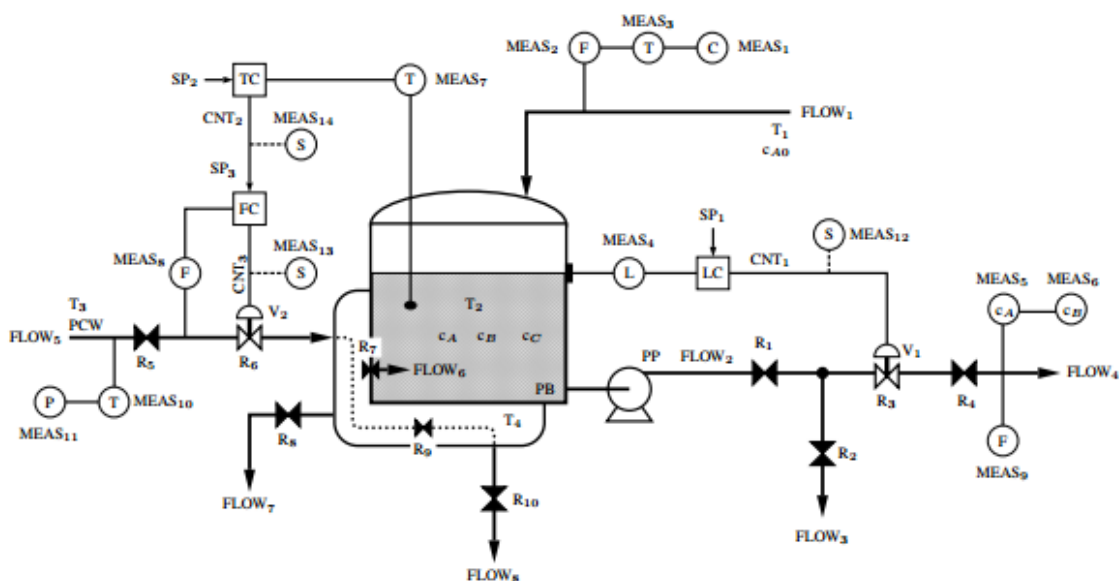
4.3.1 Descrição do simulador

O processo consiste na entrada de um reagente A com concentração inicial c_{A0} a temperatura T_1 com vazão F_1 no reator duas reações químicas acontecem em paralelo, a reação dominante ($A \rightarrow B$) é exotérmica e a segunda ($A \rightarrow C$) é endotérmica. O balanço de calor final é exotérmico, o que gera o aumento da temperatura do tanque, T_2 . Os produtos B , C e o restante do reagente A deixam o tanque com uma vazão F_2 e concentrações c_A e c_B (a concentração do subproduto c_C é ignorada).

Como o processo é predominantemente exotérmico, é necessário algum mecanismo de resfriamento do sistema. Para isso, um fluido refrigerante com vazão F_5 a temperatura T_3 entra pela jaqueta de refrigeração do reator. Refrigera o reator, que está a temperatura T_2 , maior que a temperatura do refrigerante.

O simulador permite a simulação de diferentes falhas do processo do reator. Essas falhas são adicionadas através da alteração do valor de uma variável do processo, ou através da alteração da resistência (R) para os componentes hidráulicos do processo. A Figura 16 ilustra um esquemático do reator e na Tabela 15 são mostradas as variáveis analisadas, bem como suas siglas e unidades. Todas as variáveis relacionadas a fluxo (FLOW) são abreviadas para F .

Figura 16 – O processo CSTR.



Fonte: Adaptado de Finch (1989)

Tabela 15 – Variáveis do simulador

Número	Descrição da variável	Sigla	Unidade
1	Concentração de entrada	c_{A0}	mol/m ³
2	Vazão de entrada	F_1	m ³ /s
3	Temperatura de entrada	T_1	°C
4	Nível do Reator	L	m
5	Concentração de Produto A	c_A	mol/m ³
6	Concentração de Produto B	c_B	mol/m ³
7	Temperatura do Reator	T_2	°C
8	Vazão de Refrigerante	F_5	m ³ /s
9	Vazão de Produto	F_4	m ³ /s
10	Temperatura Refrigerante	T_3	°C
11	Pressão Refrigerante	PCW	kPa
12	Sinal de Controle Nível	CNT ₁	-
13	Sinal de Controle de Vazão	CNT ₃	-
14	Setpoint Controle Vazão	CNT ₂	-
15	Setpoint Controle Nível	SP ₁	m
16	Setpoint Controle Temperatura	SP ₂	°C

Fonte: Adaptado de Finch (1989)

As falhas que podem ser aplicadas no simulador são definidas em Oyeleye (1990). A Tabela 16 lista o número de identificação de cada falha, bem como seus nomes e o parâmetro que deve ser alterado para sua simulação.

Tabela 16 – Lista de falhas do processo

Número	Nome da falha	Parâmetro afetado
1	Sem falha	-
2	Bloqueio na saída do reator	R_1
3	Bloqueio na jaqueta de resfriamento	R_9
4	Vazamento da jaqueta para o ambiente	R_8
5	Vazamento da jaqueta para o reator	R_7
6	Vazamento da bomba	R_2
7	Perda de pressão na bomba	PP
8	Mudança na superfície de troca de calor	UA
9	Fonte externa de calor	Q_{ext}
10	Energia primária de ativação	β_1
11	Energia secundária de ativação	β_2
12	Vazão anormal de reagente	F_1
13	Temperatura anormal de reagente	T_1
14	Concentração anormal de reagente	c_{A0}
15	Temperatura anormal de refrigerante	T_3
16	Pressão anormal de refrigerante	PCW
17	Pressão anormal de saída da jaqueta	JEP
18	Pressão anormal de saída do reator	REP

Fonte: Adaptado de Oyeleye (1990)

4.3.2 Análises realizadas

Para testar a aplicação da metodologia proposta para múltiplos pontos de operação, são avaliados dois tipos de situações: na primeira se tem uma quantidade finita de pontos de operação, e na segunda são analisados casos onde se tem infinitos pontos de operação. No primeiro caso, as aplicações são realizadas utilizando para classificação os dados de medição das variáveis e as contribuições das variáveis. No segundo caso, as análises são realizadas utilizando os desvios das variáveis após uma falha, os resíduos calculados com a análise externa, e a contribuição dos resíduos.

Para analisar a capacidade de generalização da metodologia, são comparadas situações onde os classificadores são treinados em um único ponto de operação, com situações onde os classificadores são treinados com dados de falhas simuladas em todos os pontos de operação.

Para efeito de comparação, são escolhidos arbitrariamente quatro diferentes pontos de operação onde serão simuladas as falhas. A Tabela 17 apresenta os valores dos *setpoints* de nível e temperatura para as regiões de operação escolhidas.

Tabela 17 – *Setpoints* para os pontos de operação escolhidos

Ponto de operação	<i>Setpoint</i> de nível	<i>Setpoint</i> de temperatura
1	2	80
2	2.1	82.5
3	2.2	85
4	2.3	87.5

Fonte: Adaptado de Oyeleye (1990)

4.3.3 Criação do modelo de variáveis externas

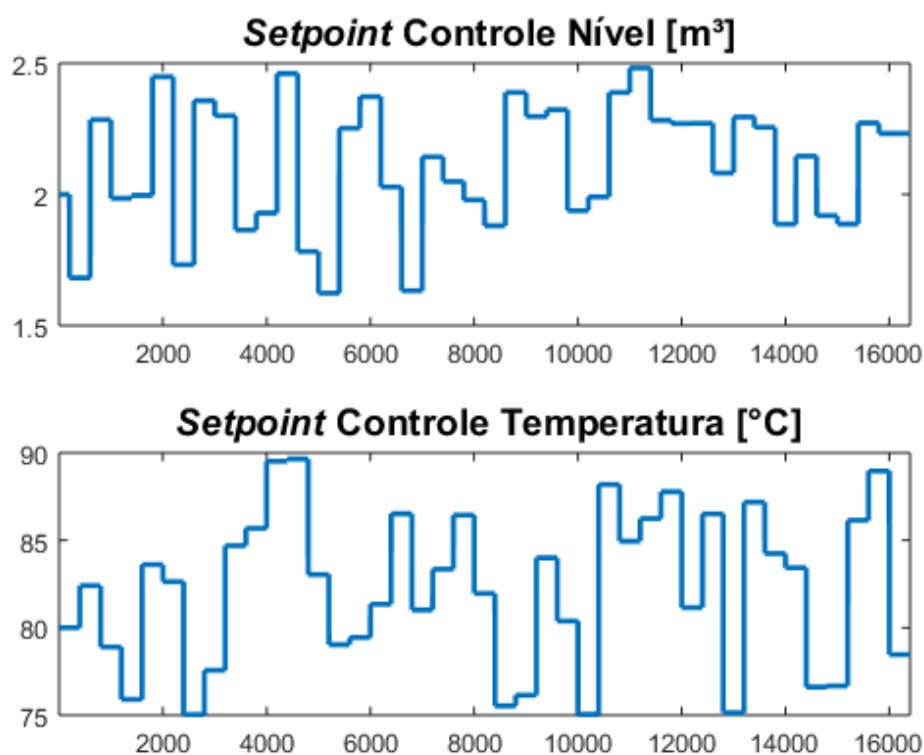
Nesse simulador as variáveis externas consideradas são os *setpoints* de controle de nível e de temperatura da planta (variáveis 15 e 16), essas variáveis são escolhidas pois elas possuem influência nos controladores, e por consequência nos valores das variáveis controladas. As demais variáveis serão divididas em dois grupos: as principais (variáveis 4, 5, 6, 7, 8 e 9), que são influenciadas diretamente pelas variáveis externas, e as variáveis independentes (variáveis 1, 2, 3, 10 e 11).

$$\begin{bmatrix} E_p \\ E_i \end{bmatrix} = \begin{bmatrix} Y_p \\ Y_i \end{bmatrix} - \begin{bmatrix} \hat{Y}_p \\ Y_{i0} \end{bmatrix} \quad (42)$$

O resíduo das variáveis principais (E_p) é calculado como a diferença entre os valores medidos (Y_p) e os valores previstos através da regressão (\hat{Y}_p) e o resíduo das variáveis independentes é calculado como a diferença entre os valores medidos (Y_i) e os valores nominais dessas variáveis (Y_{i0}).

O valor nominal dos *setpoints* de controle de nível e de temperatura são 2 metros e 80°C, respectivamente. Para construção do modelo, esses parâmetros são variados aleatoriamente entre dois limiares pré-determinados de 2,5 metros para o nível do reator e 90°C para a temperatura do mesmo. A simulação para criação do modelo é então realizada com 40 alterações de *setpoint* e a técnica de análise externa é aplicada. A Figura 17 ilustra as mudanças de *setpoint* realizadas para construção do modelo, nela ocorre uma alteração no ponto de operação da planta a cada 200 amostras.

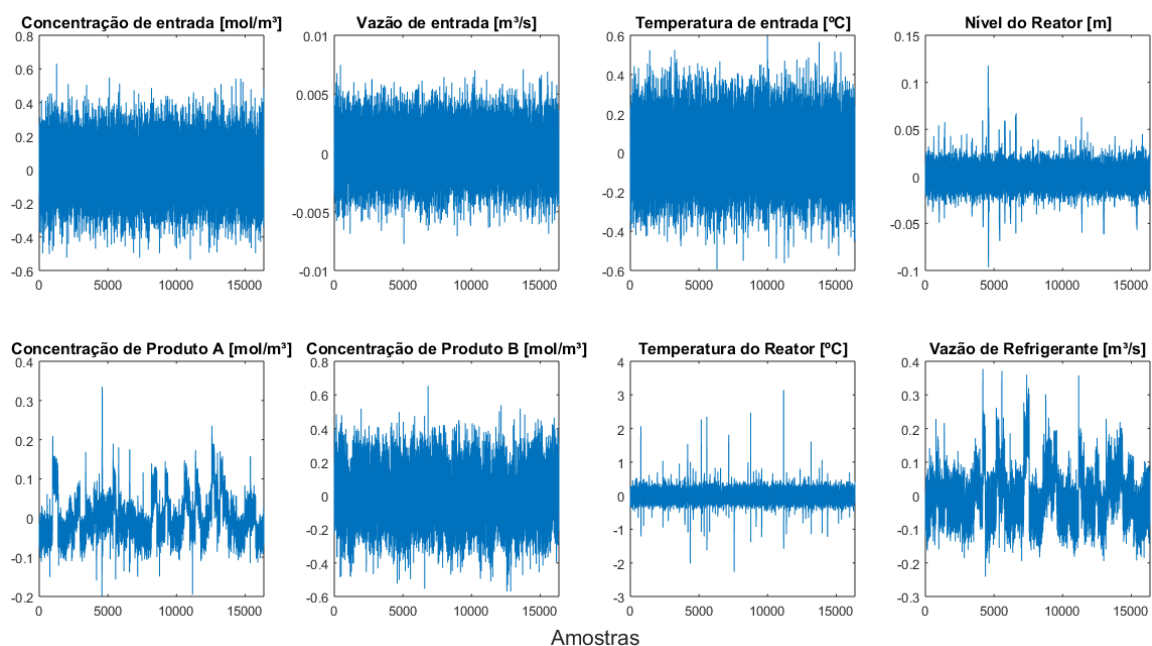
Figura 17 – Alteração dos *setpoints*.



Fonte: O próprio autor

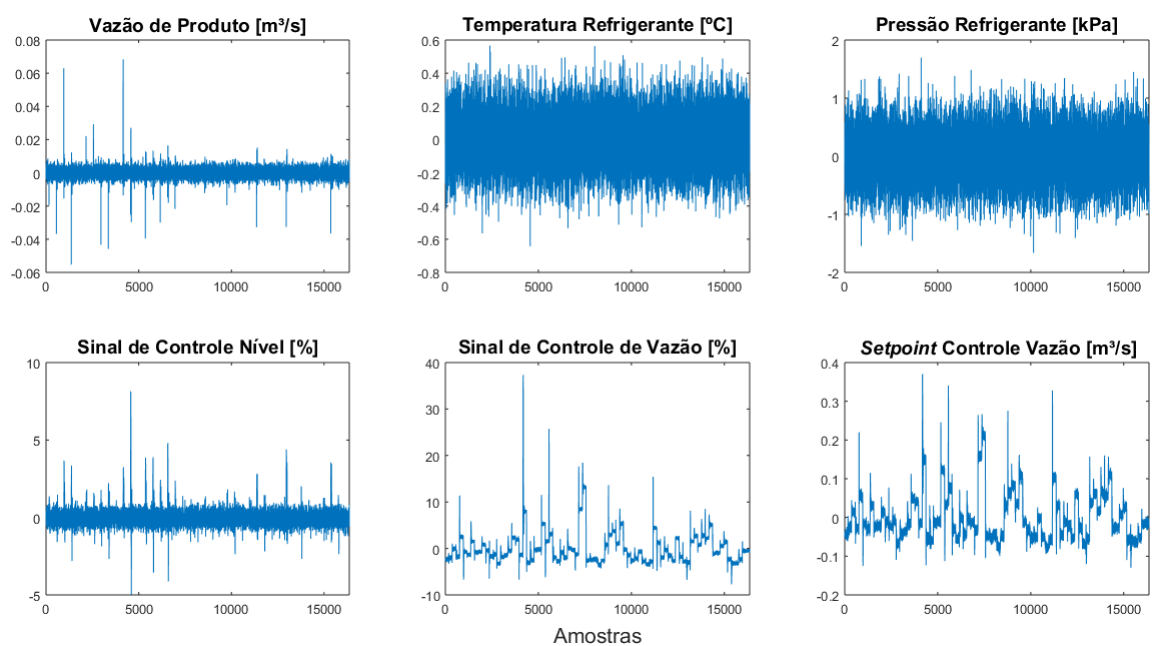
As Figuras 18 e 19 ilustram os resíduos obtidos, quando aplicada a análise externa com os parâmetros citados anteriormente, nas demais variáveis do simulador.

Figura 18 – Resíduos obtidos.



Fonte: O próprio Autor

Figura 19 – Resíduos obtidos.



Fonte: O próprio autor

Os resíduos obtidos são então utilizados para calcular o modelo PCA que será utilizado para a detecção e diagnóstico das falhas.

4.3.4 Falhas em múltiplos pontos de operação

Em cada ponto de operação são realizadas 100 simulações de cada uma das falhas, de 2 a 18, do simulador. A partir de cada uma dessas simulações é calculado o vetor de direção de falha (seção 2.1.2.4), usando os valores dos resíduos.

Em seguida são analisadas a aplicação de diferentes dados de entrada nos classificadores. Nesse estudo de caso a metodologia proposta é aplicada para quatro entradas de dados, são elas: as medições das variáveis, as contribuições das variáveis, os resíduos calculados a partir da análise externa e os desvios das variáveis.

Duas abordagens são testadas para analisar a aplicação da metodologia em diferentes pontos de operação. Na primeira abordagem, para testar se é possível aplicar um mesmo classificador em diferentes pontos de operação, a metodologia é aplicada para treinar classificadores para um ponto de operação e cada grupo de classificadores é testado em todos os pontos de operação. Em seguida, para avaliar a capacidade de generalização da metodologia para todos os pontos de operação, os classificadores locais são treinados utilizando dados de simulações de falhas nos quatro pontos de operação estudados e depois os testes são realizados para novas simulações em cada um dos pontos de operação.

4.3.5 Treinamento dos classificadores

A partir dos dados obtidos com as simulações é calculada a matriz de variáveis em falha, utilizada para a criação dos grupos. Utilizando o critério de Davies-Bouldin (seção 2.6.2) o número ótimo de grupos foi calculado, utilizando a distância de Hamming como métrica, resultando em cinco possíveis combinações de variáveis que possuíram influência na violação do limiar estatístico das falhas analisadas. A Tabela 18 mostra os resultados obtidos do agrupamento, indicando as falhas presentes em cada um dos grupos calculados, cada grupo terá dados de todas as simulações das falhas contidas no mesmo.

Tabela 18 – Resultado do agrupamento de falhas para o CSTR

Grupo	Falhas
1	2, 5, 6, 7, 18
2	3, 10, 14, 15, 16, 17
3	4, 8, 9, 10, 11, 14
4	13
5	12

Fonte: O próprio autor

Observa-se que a maioria das falhas ocorreram em apenas um grupo, os grupos possuem apenas subconjuntos de falhas. Também se observa que os grupos 4 e 5 possuem apenas uma falha. Portanto, falhas atribuídas a estes grupos serão automaticamente classificadas como falhas 13 e 12, respectivamente. Nos demais grupos é aplicado o algoritmo de classificação proposto.

Na etapa de treinamento dos modelos, 70% dos dados de cada grupo são utilizados para treinar o classificador e os demais 30% para validação. Ao realizar a divisão dos dados, é mantida a mesma proporção de dados de cada falha para ser realizado o treinamento. Deve-se ter cautela na separação dos dados para que uma mesma simulação não seja repartida, uma vez que as simulações são séries temporais e sua continuidade é importante para a classificação. Para o cálculo das características foram utilizadas 40 observações após a falha e os dados foram divididos em dois segmentos de 20 observações.

Durante o treinamento e validação dos classificadores é utilizado o conhecimento *a priori* sobre a falha ocorrida para avaliar se a aplicação da metodologia é viável para diagnóstico de falhas. Uma vez que a metodologia resulte em boas métricas de classificação são realizadas simulações de novas falhas e os classificadores treinados identificam a falha ocorrida.

Para avaliar os classificadores treinados para o diagnóstico, 100 novas simulações são geradas para cada falha, em cada ponto de operação, para atuar como conjunto de teste. Após a detecção de uma falha é extraído o vetor de direção de falha da simulação, em seguida a falha desconhecida é alocada em um grupo onde o classificador treinado realiza o diagnóstico sobre as características extraídas.

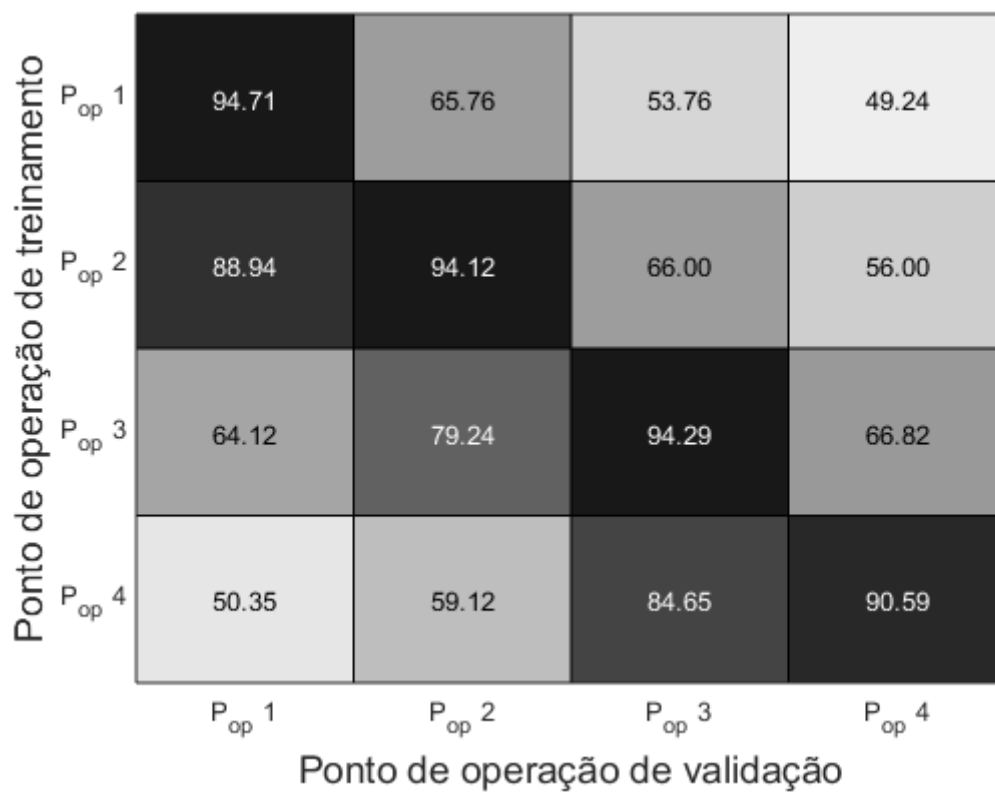
4.3.6 Classificadores treinados com as medições das variáveis

Nesse caso os classificadores são treinados utilizando as medições das variáveis (1 a 14), sem nenhum tratamento, após a detecção de uma falha. O procedimento de treinamento e validação para diferentes pontos de operação é realizado conforme apresentado na seção anterior.

A Figura 20 ilustra a matriz de confusão das acurácias dos classificadores quando são treinados classificadores independentes para cada um dos pontos de operação, onde no

eixo das ordenadas estão os pontos de operação para os quais os classificadores foram treinados e no eixo das abcissas os pontos de operação utilizados para a validação.

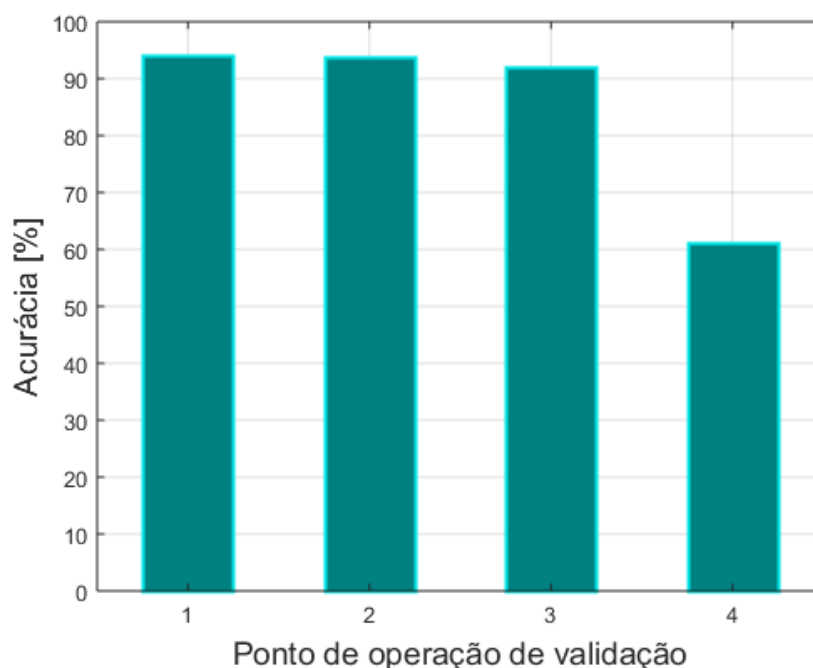
Figura 20 – Teste dos classificadores utilizando as medições das variáveis em múltiplos pontos de operação.



Fonte: O próprio autor

Na Figura 21 são apresentados os resultados obtidos são utilizados classificadores treinados com dados de todos os pontos de operação na etapa de treinamento.

Figura 21 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando as medições das variáveis.



Fonte: O próprio autor

Na Figura 20 é possível perceber que melhores acurácias se encontram na diagonal principal, uma vez que representa o ponto de operação para o qual os classificadores foram treinados, além disso, percebe-se que em pontos de operação mais afastados correram as maiores taxas de erros. Quando se conhece o ponto de operação sobre o qual a planta opera, a aplicação de classificadores específicos de cada ponto de operação demonstra melhores resultados. Porém quando o ponto de operação não é conhecido a aplicação de classificadores treinados com dados de todos os pontos de operação pode retornar bons resultados.

4.3.7 Classificadores treinados com as contribuições das variáveis

Nesse estudo os dados de contribuição das variáveis (1 a 14), conhecidos os pontos de operação, são utilizados. Esses dados são calculados a partir de modelos de PCA obtidos anteriormente para cada ponto de operação a partir da matriz característica dos modelos (seção 2.1.2.1). Para a aplicação todas as contribuições foram calculadas com base na estatística combinada (φ). A Figura 22 ilustra a matriz de confusão das acurácias dos classificadores treinados para cada ponto de operação e a Figura 23 apresenta o

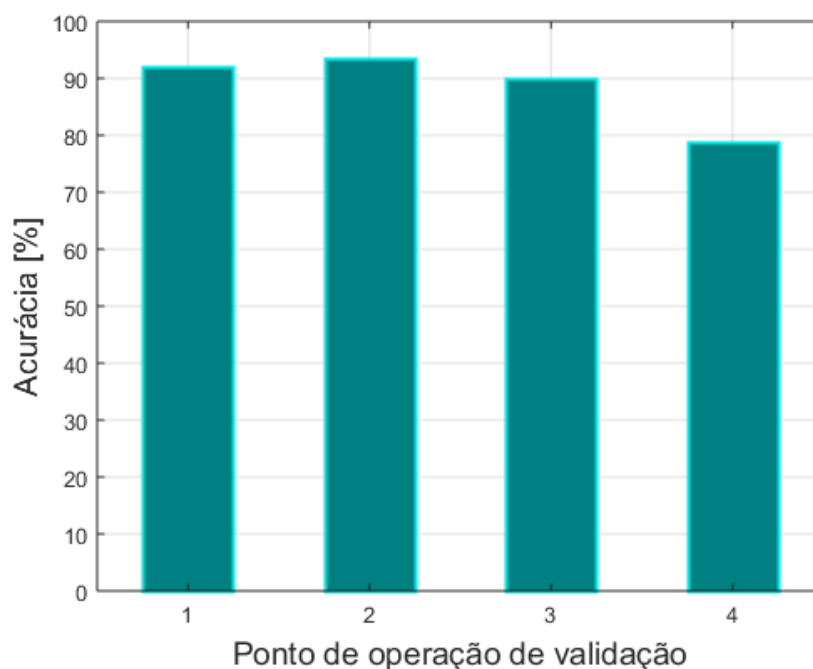
resultado quando os classificadores são treinados com dados de todos os pontos de operação.

Figura 22 – Teste dos classificadores utilizando as contribuições das variáveis em múltiplos pontos de operação.

Ponto de operação de treinamento	P _{op} 1	96.29	83.18	73.88	69.47
	P _{op} 2	87.35	93.76	87.94	75.35
	P _{op} 3	75.35	87.53	94.12	87.76
	P _{op} 4	70.76	77.71	88.06	93.12
		P _{op} 1	P _{op} 2	P _{op} 3	P _{op} 4
Ponto de operação de validação					

Fonte: O próprio autor

Figura 23 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando as contribuições das variáveis.



Fonte: O próprio autor

A alta acurácia dos resultados das aplicações quando conhecidos os pontos de operação motivaram uso de dados que independam do ponto de operação, para isso foram explorados o uso de contribuições, desvios, resíduos e contribuições dos resíduos obtidos com a análise externa.

4.3.8 Classificadores treinados com os desvios das variáveis

Nesse caso, a influência do ponto de operação é removida através da remoção da média dos dados anterior a detecção de uma falha. São calculadas as médias das variáveis, em um determinado intervalo anterior a detecção da falha, que é removida dos dados em falha, resultando nos desvios das variáveis após a ocorrência de uma falha.

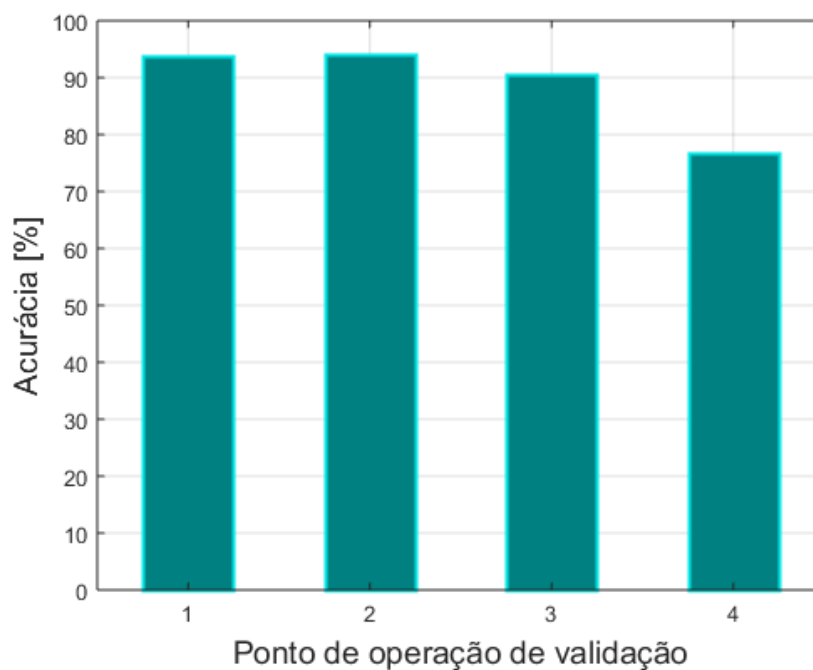
Os desvios das variáveis 1 a 14 são aplicados na metodologia e novamente dois grupos de classificadores são treinados, a Figura 24 apresenta a acurácias dos testes quando classificadores específicos de cara ponto de operação são aplicados aos demais. A Figura 25 apresenta o resultado obtido quando os classificadores são treinados com dados de falhas de todos os pontos de operação.

Figura 24 – Teste dos classificadores utilizando os desvios das variáveis em múltiplos pontos de operação.

Ponto de operação de treinamento	P _{op} 1	94.29	82.65	76.59	75.65
	P _{op} 2	92.71	93.82	83.24	76.53
	P _{op} 3	82.29	89.94	94.18	82.41
	P _{op} 4	77.94	83.76	90.00	90.82
		P _{op} 1	P _{op} 2	P _{op} 3	P _{op} 4
Ponto de operação de validação					

Fonte: O próprio autor

Figura 25 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando os desvios das variáveis.



Fonte: O próprio autor

Mesmo com a aplicação dos desvios, a influência do ponto de operação no comportamento das variáveis após uma falha não é totalmente removida. Novamente percebe-se que, quando o ponto de operação é conhecido é possível obter diagnósticos com uma maior qualidade. Porém, quando comparados os resultados obtidos nas Figuras 20 e 24 é possível perceber um ganho na acurácia das falhas testadas no ponto de operação 4, indicando que é possível obter bons resultados quando não se tem conhecimento prévio do ponto de operação da planta. Com isso são realizados testes com dados de resíduos e contribuição dos resíduos, que são menos afetados pela alteração do ponto de operação, com o objetivo de encontrar um grupo de classificadores capaz de diagnosticar corretamente uma falha independente do ponto de operação.

4.3.9 Classificadores treinados com os resíduos

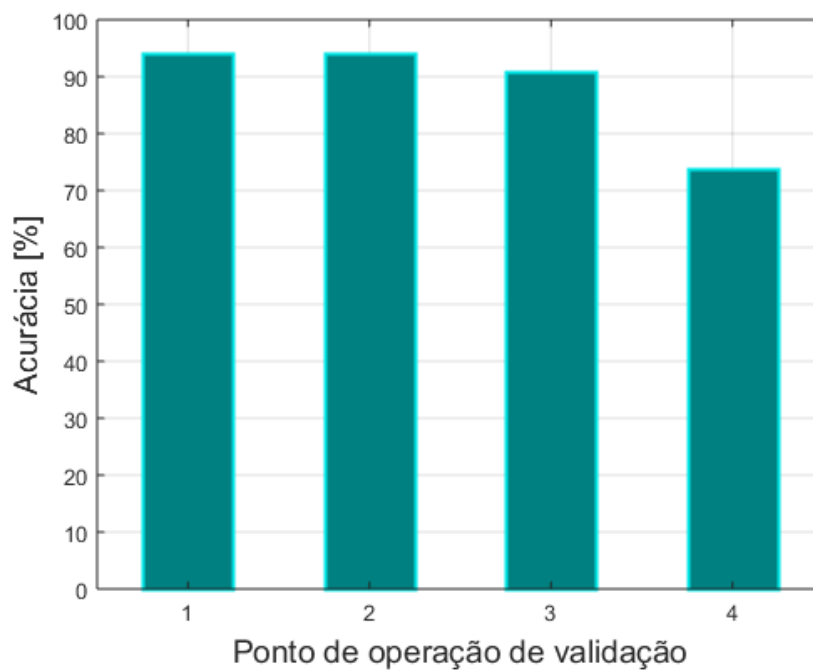
Nessa aplicação é utilizado o modelo de regressão obtido com a análise externa (seção 4.3.3) para calcular os resíduos das variáveis principais (1 a 14) em relação ao modelo obtido com as variáveis externas (15 e 16). Com a ocorrência de um evento de falha a medição das variáveis se afastará dos valores previstos pelo modelo, aumentando assim o resíduo. Supõe-se que os resíduos obtidos de uma mesma falha se comportem de forma semelhante em diferentes pontos de operação. Os valores dos resíduos após a detecção das falhas são utilizados para treinar os classificadores com as mesmas proporções de dados de treinamento e classificação apresentadas na seção 4.3.5.

Figura 26 – Teste dos classificadores com resíduos em múltiplos pontos de operação

Ponto de operação de treinamento	P _{op} 1	P _{op} 2	P _{op} 3	P _{op} 4
	94.47	82.18	76.00	71.65
	92.06	94.06	83.12	73.29
	82.41	89.65	94.29	83.06
P _{op} 4	77.59	82.00	88.82	91.00
Ponto de operação de validação				

Fonte: O próprio autor

Figura 27 – Teste dos classificadores treinados com dados de todos os pontos de operação utilizando os resíduos.



Fonte: O próprio autor

No caso estudado, a aplicação de classificadores treinados com falhas em todos os pontos de operação obteve um resultado médio superior ao da aplicação de classificadores treinados por ponto de operação, uma vez que as acurácias médias foram 88,18% e 84,72%, respectivamente.

A aplicação de classificadores quando se tem conhecimento do ponto de operação da planta obteve resultados melhores (diagonal da matriz de confusão), porém o uso de um único classificador obteve uma menor acurácia para o ponto de operação 4 quando comparado ao resultado utilizando os desvios das variáveis principais.

4.3.10 Classificadores treinados com as contribuições dos resíduos

Nessa aplicação os dados de contribuição (seção 2.2.1) dos resíduos das variáveis principais (1 a 14) são utilizados, esses dados são calculados utilizando o modelo PCA obtido a partir dos resíduos calculados na seção 4.3.3. Para a aplicação todas as contribuições foram calculadas com base na estatística combinada (φ).

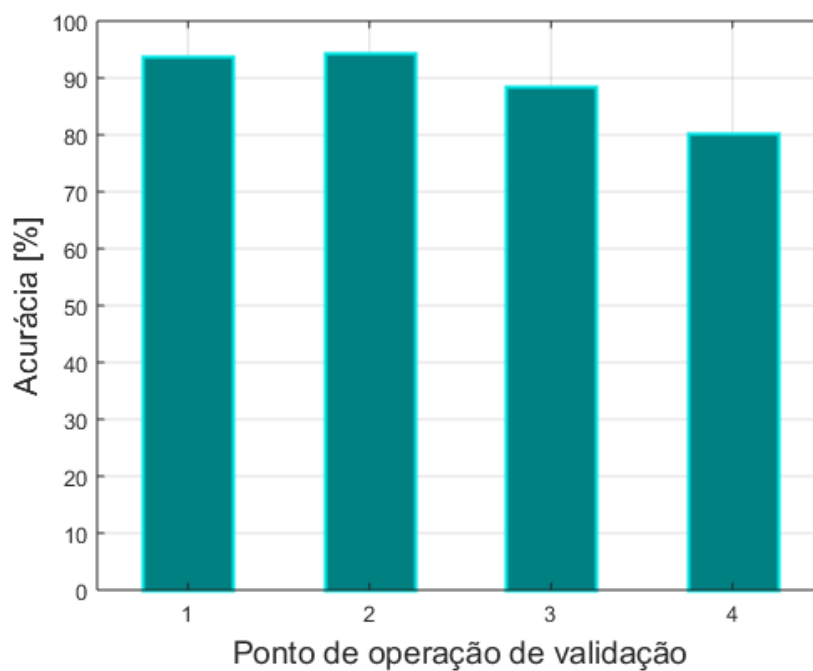
A Figura 28 ilustra a matriz de confusão das acurácias do teste dos classificadores treinados para diferentes pontos de operação, Figura 29 apresenta o resultado do teste quando os classificadores são treinados com dados de falhas de todos os pontos de operação.

Figura 28 – Teste dos classificadores utilizando as contribuições dos resíduos em múltiplos pontos de operação

Ponto de operação de treinamento	P_{op1}	94.76	83.47	76.71	71.82
	P_{op2}	94.82	94.24	85.76	71.94
	P_{op3}	87.35	88.00	94.29	82.65
	P_{op4}	81.59	86.47	87.59	88.88
		P_{op1}	P_{op2}	P_{op3}	P_{op4}
Ponto de operação de validação					

Fonte: O próprio autor

Figura 29 – Teste dos classificadores treinados com dados de falhas de todos os pontos de operação utilizando as contribuições dos resíduos.



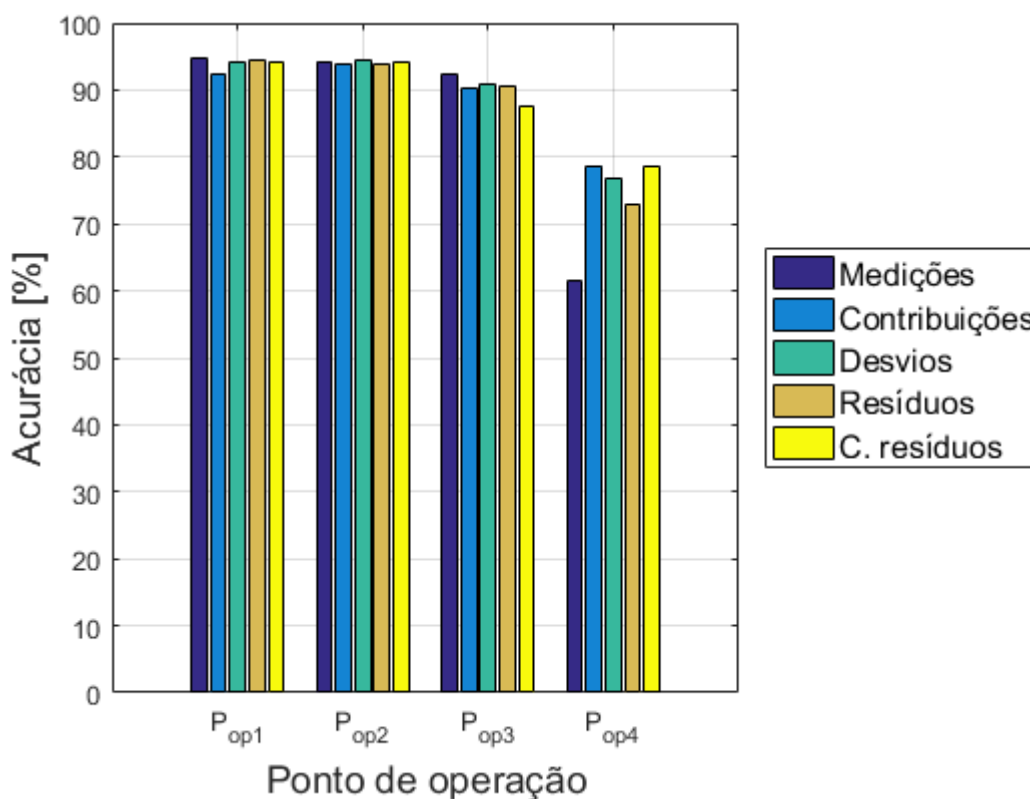
Fonte: O próprio autor

Verifica-se que o uso de contribuições forneceu os melhores resultados para diagnóstico de falhas em múltiplos pontos de operação, quando não se tem conhecimento sobre qual ponto de operação da planta opera. Porém, os melhores resultados de acurácia dos diagnósticos ainda ocorrem quando se conhece o ponto de operação que a planta está operando no momento da falha.

Quando não se conhece o ponto de operação o uso das contribuições com classificadores treinados com todos os pontos de operação resultou numa acurácia média de 91,12%, em comparação a acurácia média de 85,64% com o uso de classificadores específicos de cada ponto de operação. Isso indica que, em plantas com infinitos pontos de operação o uso de um classificador treinado com a contribuição dos resíduos oferece diagnósticos com elevada acurácia, independente do ponto de operação.

Os resultados obtidos com a aplicação da metodologia para diferentes tipos de entradas analisados são comparados na Figura 30, nela os classificadores são treinados e aplicados em dados de todos os pontos de operação.

Figura 30 – Comparação das acurácias obtidas para diferentes sinais e pontos de operação



Fonte: O próprio autor

É possível perceber que as menores acurácias são obtidas quando são testados os dados do ponto de operação 4, percebe-se também que para esse ponto de operação as medições das variáveis têm a pior acurácia entre os sinais comparados e, quando utilizados os sinais de contribuição são obtidas as melhores acurácias. Assim quando se conhece os pontos de operação da planta ou quando é utilizada a análise externa, o uso de contribuições oferece bons resultados para diagnóstico de falhas.

5 Conclusão

Neste trabalho foi apresentada uma metodologia baseada em dados para o diagnóstico de falhas em processos industriais, utilizando métodos multivariados de monitoramento de processos, agrupamento de falhas com influência em variáveis semelhantes, extração de características e classificadores locais para identificar as séries temporais resultantes das falhas e por fim classificá-las.

A abordagem proposta foi aplicada em *benchmarks* da literatura e obteve bons resultados quando um grupo de falhas bem definidas são usadas e o ponto de operação da planta é conhecido. Para plantas que operem em mais de um ponto de operação foram apresentadas e comparadas diversas metodologias de obtenção de dados para aplicação na metodologia, demonstrando que aplicação de técnicas de cálculo de resíduos, como a análise externa, podem ser aplicadas para essas plantas, e uma vez que a metodologia seja aplicada e os classificadores treinados com dados de falha em alguns pontos de operação, se torna possível utilizar um único grupo de classificadores para diagnosticar falhas, com elevada acurácia, em todos os pontos de operação.

As limitações na metodologia aparecem em casos onde diferentes falhas se manifestam de forma semelhante no espaço de características extraídas, mas, mesmo que os classificadores treinados não sejam capazes de classificar corretamente todas as falhas, o agrupamento pode ajudar a diminuir as categorias de falhas em potencial quando uma nova amostra de dados precisa ser diagnosticada. O conjunto de características estatísticas aqui utilizado é um modelo relativamente simples e pode não ser apropriado para distinguir certos casos.

Trabalhos futuros para aprimorar a metodologia incluem a extração de diferentes características, como uso da transformada *wavelet*, o uso de técnicas de *deep learning* (aprendizado profundo) para extração e seleção de características obtidas dos sinais analisados. Também podem ser estudadas outras métricas e índices para o cálculo do número ideal de grupo, uso de outras metodologias de agrupamento que permitam que o usuário separe previamente classes de difícil classificação e o uso técnicas de afinação para obter conjuntos otimizados de parâmetros do classificador *random forest*, ou mesmo o uso de outros classificadores. Por fim a realização de testes *online* da metodologia proposta.

6 Referências

- ABDI, H; WILLIAMS, L J. Principal component analysis. **Wiley interdisciplinary reviews: computational statistics**, v. 2, n. 4, p. 433-459, 2010.
- AGRAWAL, R.; FALOUTSOS, C.; SWAMI, A. Efficient similarity search in sequence databases. v. 8958546, p. 69–84, 1993.
- ALCALA, C. F.; QIN, S. J. Reconstruction-based Contribution for Process Monitoring. **Automatica**, v. 45, n. 7, p. 1593–1600, 2009.
- ALDRICH, C.; AURET, L. **Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods**. Springer, 2016.
- BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.
- BREIMAN, L.; CUTLER, A. An introduction to random forest for beginners. 1. ed. Califórnia, Estados Unidos: Salford Systems, 2014. 71 p. 24, 26, 27
- CHEN, W. **Multivariate Statistical Process Control in Industrial Plants**. Tese de doutorado. Delft University of Technology, 2005.
- CHIANG, L. H.; RUSSEL, E. L.; BRAATZ, R. D. **Fault detection and diagnosis in industrial systems**. Springer Science & Business Media, 2000.
- DANG, T. N.; WILKINSON, L. TimeExplorer: Similarity search time Series by their signatures. **Proc. Intl. Symp. Visual Computing (ISVC)**, p. 280–289, 2013.
- D'ANGELO, M. F.S.V. et al. A new fault classification approach applied to Tennessee Eastman benchmark process. **Applied Soft Computing**, v. 49, p. 676-686, 2016.
- DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, n. 2, p. 224-227, 1979.
- DIETTERICH, T. G. Approximate statistical tests for comparing supervised classification learning algorithms. **Neural Computation**, v. 10, n. 7, p. 1895–1923, 1998.
- DOU, D.; ZHOU, S. Comparison of four direct classification methods for intelligent fault diagnosis of rotating machinery. **Applied Soft Computing**, v. 46, p. 459-468, 2016.
- DOWNS, J. J.; VOGEL, E. F. A plant-wide industrial process control problem. **Computers and Chemical Engineering**, v. 17, n. 3, p. 245–255, 1993.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification** New York: John Wiley, Section, 2001.
- ESLAMLOUEYAN, R. Designing a hierarchical neural network based on fuzzy clustering for fault diagnosis of the Tennessee Eastman process. **Applied Soft Computing**, v. 11, n. 1, p. 1407-1415, 2011.
- FINCH, F. E. **Automated fault diagnosis of chemical process plants using model-based reasoning**. Tese de doutorado. University of California, 1989.
- GAO, X.; HOU, J. An improved SVM integrated GS-PCA fault diagnosis approach of Tennessee Eastman process. **Neurocomputing**, v. 174, p. 906-911, 2016.

- HOTELLING, H. The economics of public recreation. **The Prewitt Report**, 1947.
- ISERMANN, R. Supervision, fault-detection and fault-diagnosis methods - an introduction. **Control Engineering Practice**, v. 5, n. 5, p. 639–652, 1997.
- JACKSON, J. E.; MUDHOLKAR, G. S. Control procedures for residuals associated with principal component analysis. **Technometrics**, v. 21, n. 3, p. 341-349, 1979.
- JAIN, A. K.; DUBES, R. C.; CHEN, C. C. Bootstrap Techniques for Error Estimation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. PAMI-9, n. 5, p. 628–633, 1987.
- JING, C.; HOU, J. SVM and PCA based fault classification approaches for complicated industrial process. **Neurocomputing**, v. 167, p. 636-642, 2015.
- KANO, M. et al. Evolution of multivariate statistical process control: application of independent component analysis and external analysis. **Computers & Chemical Engineering**, v. 28, n. 6-7, p. 1157-1166, 2004.
- KASSAMBARA, A. Practical Guide To Cluster Analysis in R. 2015.
- KOURTI, T. Process analysis and abnormal situation detection: from theory to practice. **IEEE control systems**, v. 22, n. 5, p. 10-25, 2002.
- KRONBAUER, A. M.; FONTOURA, L. M.; WINCK, A. T. Um estudo sobre processos para avaliação de algoritmos de agrupamento de dados. **Revista ComInG-Communications and Innovations Gazette**, v. 1, n. 1, p. 34–45, 2016.
- KU, W.; STORER, R. H.; GEORGAKIS, Christos. Disturbance detection and isolation by dynamic principal component analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 30, n. 1, p. 179-196, 1995.
- KULKARNI, A.; JAYARAMAN, V. K.; KULKARNI, B. D. Knowledge incorporated support vector machines to detect faults in Tennessee Eastman Process. **Computers and Chemical Engineering**, v. 29, n. 10, p. 2128–2133, 2005.
- LAU, C. K. et al. Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS. **Chemometrics and Intelligent Laboratory Systems**, v. 120, p. 1-14, 2013.
- LI, C. et al. Observer-biased bearing condition monitoring: From fault detection to multi-fault classification. **Engineering Applications of Artificial Intelligence**, v. 50, p. 287-301, 2016.
- LI, D. et al. Time series classification with discrete wavelet transformed data: Insights from an empirical study. In: **The 28th International Conference on Software Engineering and Knowledge Engineering (SEKE 2016)**. 2016.
- LI, G.; QIN, S. J.; CHAI, T. Multi-directional reconstruction-based contributions for root-cause diagnosis of dynamic processes. In: **American Control Conference (ACC), 2014**. IEEE, 2014. p. 3500-3505.
- LIU, J.; CHEN, D. S. Fault isolation using modified contribution plots. **Computers and Chemical Engineering**, v. 61, p. 9–19, 2014.
- MAATEN, L. V. D.; HINTON, G. Visualizing Data using t-SNE. **Journal of Machine**

Learning Research 1, v. 620, n. 1, p. 267–84, 2008.

MACGREGOR, J. F. et al. Process monitoring and diagnosis by multiblock PLS methods. **AIChE Journal**, v. 40, n. 5, p. 826–838, 1994.

MAURYA, M. R.; RENGASWAMY, R.; VENKATASUBRAMANIAN, V. Fault diagnosis using dynamic trend analysis: A review and recent developments. **Engineering Applications of Artificial Intelligence**, v. 20, n. 2, p. 133–146, 2007.

MÖRCHEN, F. Time series feature extraction for data mining using DWT and DFT. **Technical Report, No. 33, Department of Mathematics and Computer Science, University of Marburg, Germany**, p. 1–31, 2003.

NOROUZI, M. et al. Hamming distance metric learning. **Nips**, p. 1–9, 2012.

OYELEYE, O. O. **Qualitative modeling of continuous chemical processes and applications to fault diagnosis**. Tese de doutorado. Massachusetts Institute of Technology, 1990.

PEARSON, K. Principal components analysis. **The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science**, v. 6, n. 2, p. 559, 1901.

QIN, S. J. Statistical process monitoring: basics and beyond. **Journal of Chemometrics**, v. 17, n. 8–9, p. 480–502, 2003.

RAGAB, A. et al. Fault detection and diagnosis in the Tennessee Eastman Process using interpretable knowledge discovery. **Proceedings - Annual Reliability and Maintainability Symposium**, 2017.

RATO, T. J.; REIS, M. S. Defining the structure of DPCA models and its impact on process monitoring and prediction activities. **Chemometrics and Intelligent Laboratory Systems**, v. 125, p. 7486, 2013.

RENGASWAMY, R.; VENKATASUBRAMANIAN, V. A fast training neural network and its updation for incipient fault detection and diagnosis. **Computers and Chemical Engineering**, v. 24, n. 2–7, p. 431–437, 2000.

SAMMON, J. W. A Nonlinear Mapping for Data Structure Analysis. **IEEE Transactions on Computers**, v. C-18, n. 5, p. 401–409, 1969.

SHEWHART, W. A. **Economic control of manufactured product**. van Nostrand, 1931.

SOUSA, D. P.; BARRETO, G. A.; MEDEIROS, C. M. S. **Efficient Selection of Data Samples for Fault Classification by the Clustering of the SOM**. Congresso Brasileiro de Inteligência Computacional. Niterói, RJ: 2017

VAN DEN KERKHOF, P. et al. Analysis of smearing-out in contribution plot based fault isolation for statistical process control. **Chemical Engineering Science**, v. 104, p. 285–293, 2013.

VENKATASUBRAMANIAN, V. et al. A review of process fault detection and diagnosis: Part I: Qualitative models and search strategies. **Computers and Chemical Engineering** v. 27, p. 313–326, 2003.

VENKATASUBRAMANIAN, V. et al. A review of process fault detection and

diagnosis: Part III: Process history based methods. **Computers & Chemical Engineering**, v. 27, p. 293–311, 2003.

VERRON, S.; TIPLICA, T.; KOBI, A. Fault diagnosis with bayesian networks: Application to the Tennessee Eastman process. In: **Industrial Technology, 2006. ICIT 2006. IEEE International Conference on**. IEEE, 2006. p. 98-103.

WESTERHUIS, J. A.; GURDEN, S. P.; SMILDE, A. K. Generalized contribution plots in multivariate statistical process monitoring. **Chemometrics and intelligent laboratory systems**, v. 51, n. 1, p. 95-114, 2000.

YUE, H. H.; QIN, S. J. Reconstruction-Based Fault Identification Using a Combined Index. **Industrial & Engineering Chemistry Research**, v. 40, n. 20, p. 4403–4414, 2001.

ZHANG, J. Improved on-line process fault diagnosis through information fusion in multiple neural networks. **Computers & Chemical Engineering**, v. 30, n. 3, p. 558-571, 2006.